

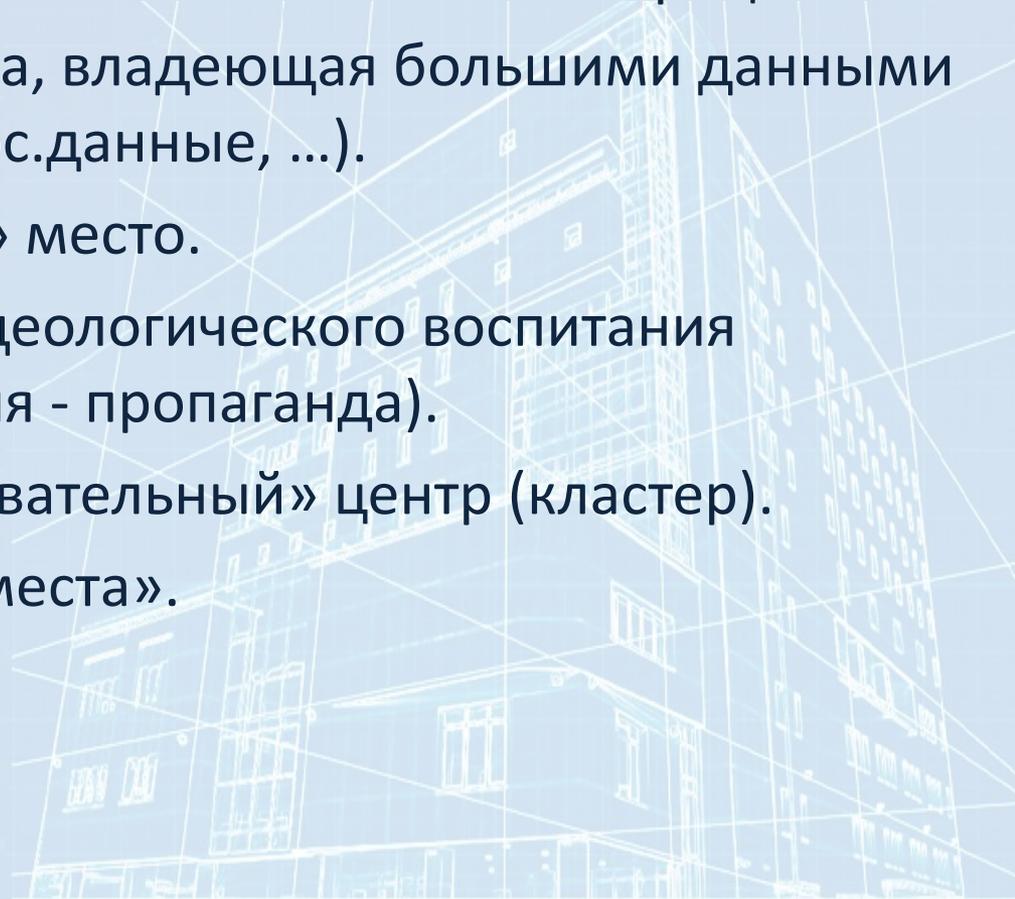


# Научно-техническая библиотека в современных условиях

**Карауш Александр Сергеевич,**

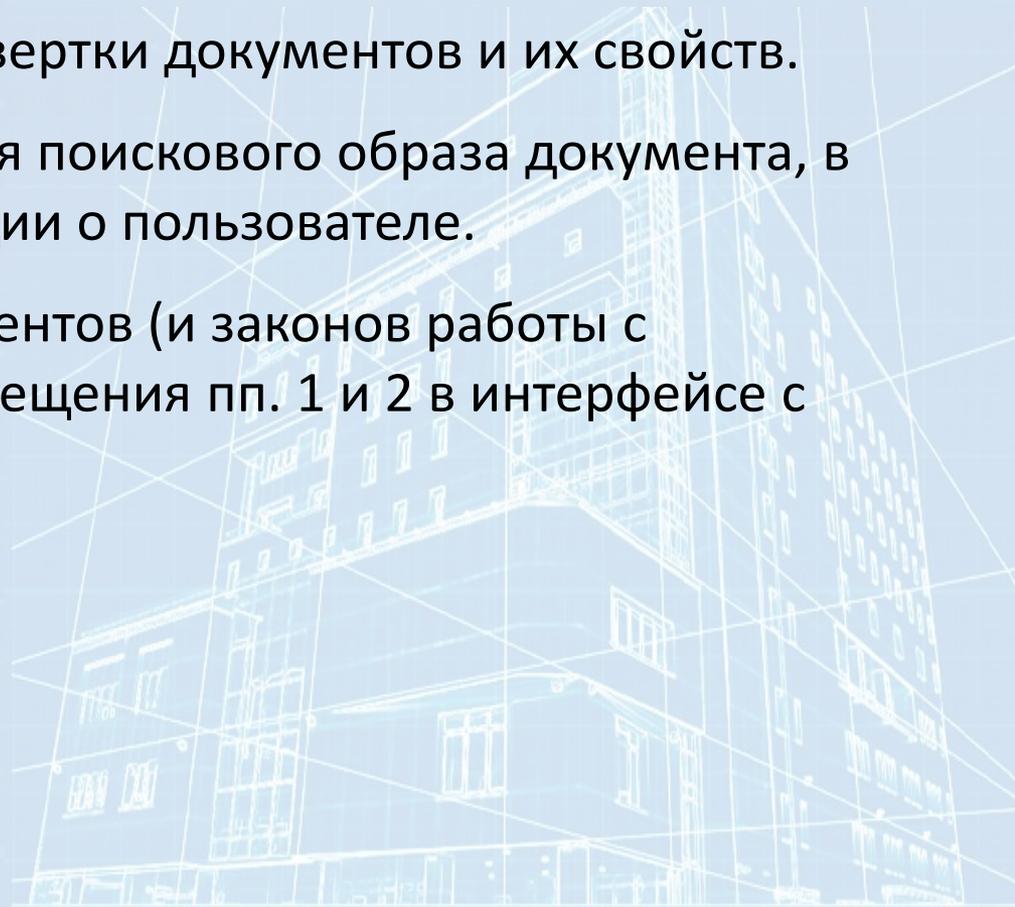
Генеральный директор,  
Государственная публичная  
научно-техническая библиотека России, к.т.н,  
Вице-президент, Национальная библиотечная  
ассоциация «Библиотеки будущего»

1. Библиотека – структурное подразделение, обслуживающее основные технологические процессы.
2. Библиотека – структура, владеющая большими данными (фонд, статистика, перс.данные, ...).
3. Библиотека – «третье» место.
4. Библиотека – звено идеологического воспитания (онтология – идеология - пропаганда).
5. Библиотека – «образовательный» центр (кластер).
6. Библиотека – «гений места».
7. Библиотека – ...



## Задачи библиотечного дела: (некоторые)

1. Создание (получение) свертки документов и их свойств.
2. Создание и актуализация поискового образа документа, в т.ч. на основе информации о пользователе.
3. Использование инструментов (и законов работы с информацией) для совмещения пп. 1 и 2 в интерфейсе с пользователем.



# Меняются институты, поддерживающие циклы обращения знаний

На функции кандидаткой «клеточки» Новой промышленной революции претендуют модульные технологические «платформы» — в том числе, с открытой архитектурой.

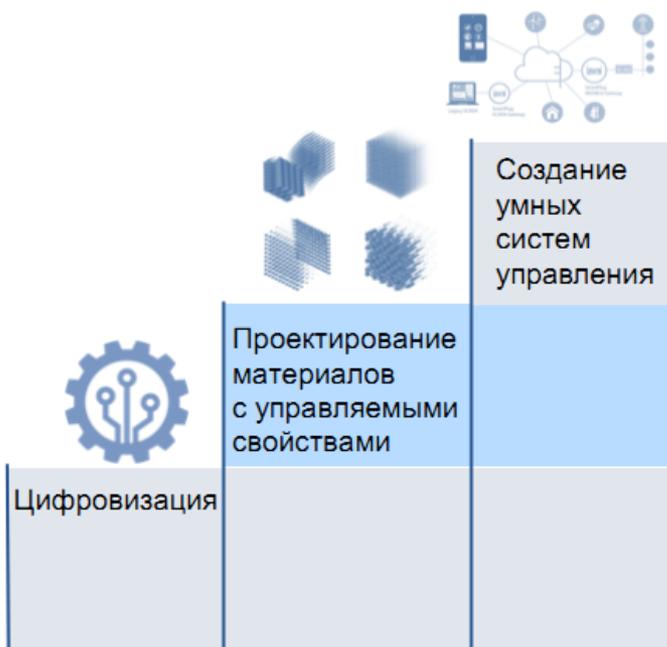


# Изменение социальной организации общества



# Изменения «железных технологий» сегодня идут по трём основным технологическим коридорам

## Направления развития



McKinsey & Company так оценивает ежегодный рост производительности труда от освоения новых технологий на производстве в период очередной промышленной революции:

Технологии	Рост производительности труда в год <sup>2</sup>
Паровой двигатель	1850-1910 0,30%
Ранние роботы	1993-2007 0,40%
IT	1995-2005 0,60%
Роботы, ИИ, машинное обучение	2015-2065 0,8-1,4%

**Сложность**

<sup>1</sup> McKinsey & Company, Industry 4.0 How to navigate digitization of the manufacturing sector, 2015 <sup>2</sup> McKinsey & Company, <http://www.mckinsey.com/global-themes/digital-disruption/harnessing-automation-for-a-future-that-works>

1. Объединение основных данных в руках государства (Гостех, НЭБ, ссылки, поиск, лингвистика ...)
2. Унификация всех технологических процессов за счет усложнения и укрупнения (оборудование, стандарты, регламенты, список разрешенных систем, оборудования и пр.)
3. Минимизация внешних к библиотеке ИТ процессов (ИТ кадры, обслуживание, безопасность...)
4. Пузырь рекомендательных систем (зона комфорта).
5. Применение исключительно внешних алгоритмов и правил обработки любых данных, что имеются в библиотеках (сводные каталоги, БД читателей, БД посещений, списки оборудования, площадей, ресурсов...)
6. Стремление участвовать в процессах построения цифрового суверенитета государства.

1. Люди с удовольствием сваливают свои проблемы на роботов (ИИ).
2. В большом потоке данные человек будет бессилён ориентироваться. Поэтому фальсификация всего, что невозможно быстро проверить (документы, законы, история, идеологические составляющие, персоны, фото, события, аккаунты, комменты, сайты целиком, дипломы, научные работы и пр.).
3. Изменение состава профессионального «околокнижного» сообщества.
4. Риск цифровой колонизации «по полному варианту».
5. ИИ – это не проект, а целая индустрия, часть которой не видна никому, т.к. логика построения инфраструктуры (оборудование, суперкорпорации, частные большие проекты) нигде не описаны.

## Свойства конечного потребителя информации библиотек:

- Ни один из вновь улучшенных поисковых инструментов работы с информацией не приводил очередное поколение пользователей к повышению качества их поисковых возможностей (способностей).
- Все инструменты с задачей измерения (и метрики) пользователей и их перспектив, а также их оценки были взломаны для получения измеримых сиюминутных выгод.
- Любая алгоритмизация деятельности приводит к появлению инструментов (СРТ) для увеличения производительности поискового труда путём упрощения его составляющих.

# Инструменты оптимизации (ИИ) мыслительной деятельности для принятия решений

- Типизация результатов, полученных на основе алгоритмов с обучением и подкреплением (на безрыбье и рак-рыба).
- Невозможность уникальных решений, как и уникальных запросов (примитивизация промптов).
- Потеря осознанности принятого решения (в поисковике так было, система так показала, ИИ так рекомендовал, поскольку вероятность совпадения 90%, то значит – истина, ...)

В 2019 году инженеры Microsoft исследовали болезнь Morbus Google. Смысл в том, что если вы не обладаете общими знаниями по проблеме, то и поиск в сети по вам ничего не даст.

Необходимо получить начальное образование по данной теме, чтобы уже потом детально разбираться и уметь грамотно сортировать информацию. Но если нет начальных знаний, то поисковик вам не поможет, а только запутает вас.

Онлайн образование осваивается хуже, чем образование через книгу. Книга линейна и шаг за шагом объясняет предмет, вводит в курс дела, раскрывает знания, на основе которых уже в дальнейшем происходит понимание темы и картины знаний, идёт осмысленное формирование картины окружающего мира, что приводит к уменьшению затрат бесцельно проведённого времени около поисковика.

Чем меньше предварительных знаний, тем меньше может помочь инструмент, который работает с информацией (поисковик, ИИ и пр.).

# Парадигма Data-driven Science меняет...

## Система научных коммуникаций

### Публикация

- название, авторы
- аннотация
- ключевые слова

Текст

Исследователь

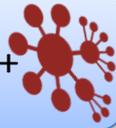


Объекты научной коммуникации

Субъект научной коммуникации

### Публикация

- название, авторы
- аннотация
- ключевые слова

Текст + 

методы  
объекты  
свойства  
эффекты  
модели

Цифровой  
помощник

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



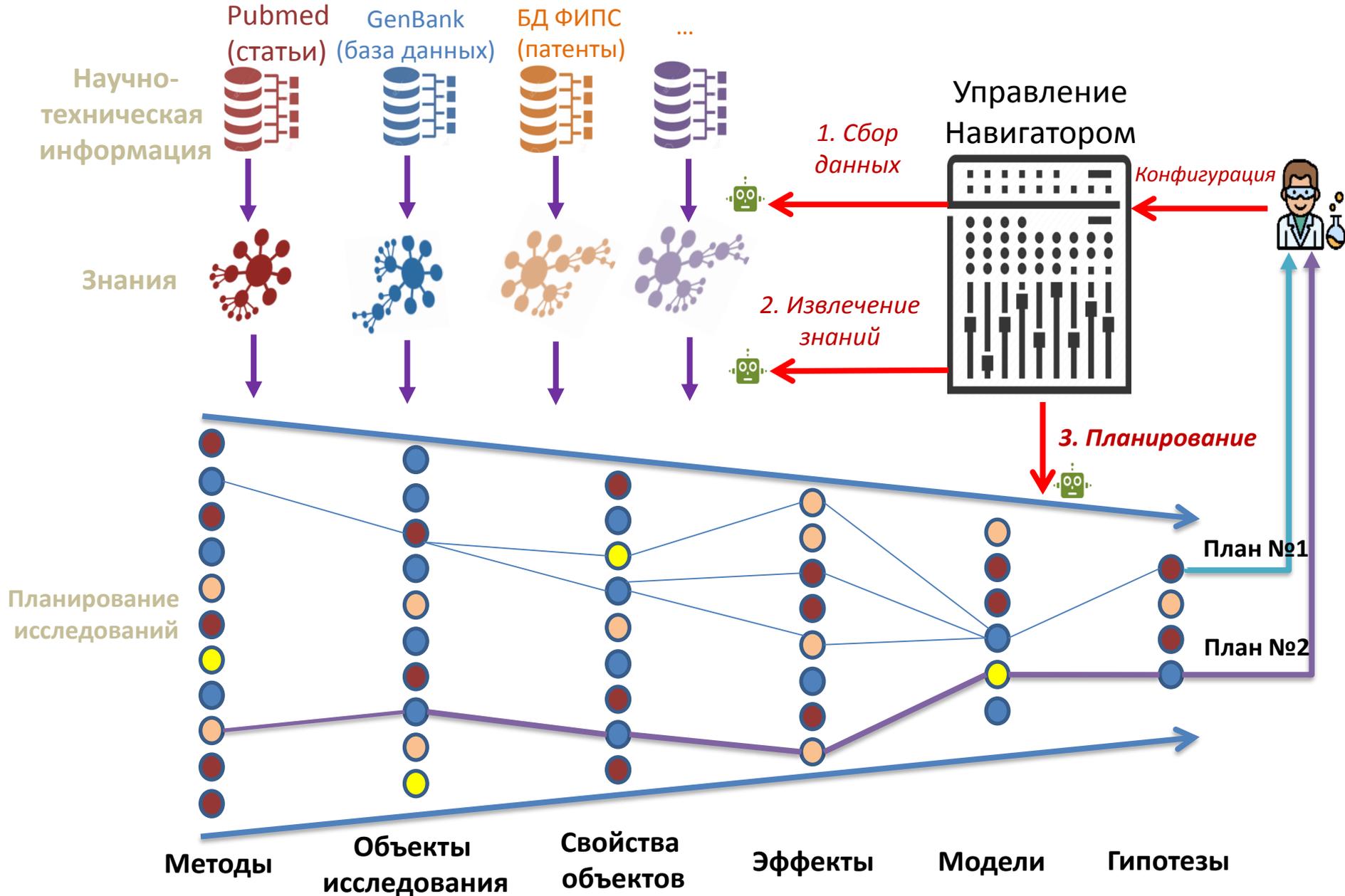
## Инструменты жизненного цикла исследований



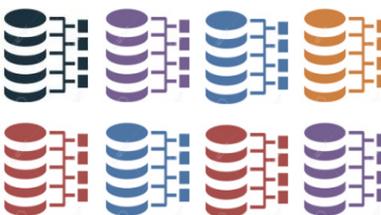
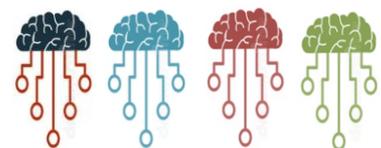
Научно-информационная инфраструктура

Было  
Стало

# Пример из будущего: Навигатор для планирования исследований



# Принцип приращения ценности научной информации



Приращение 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

# 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

## Задачи для специалиста библиотеки

- Не всегда верить «железке» во всём.
- Понимать пределы технологий, в т.ч. пределы возможностей ИИ.
- Постоянно оттачивать навыки собственноручного использования инструментов с ИИ, чтобы не смотреть с позиции «розовых очков» в каждую новую технологию и инструмент.
- Не терять навыки работы с живыми читателями, даже если между вами ИИ.
- Всегда включать логику социальной нужности для общества в будущем, важности личного общения, отвечать на вопросы потери индивидуальности во всём.

## Вывод

- ИИ (равно / не равно) «черный ящик»?
- ИИ – инструмент или актер?
- Всё, что делается последние 10 лет и не понимается в плане работы алгоритмов, тоже можно назвать ИИ в понимании «черного ящика» для конкретного, глубоко в алгоритмах не разбирающегося, человека.
- Если ИИ показывает более высокую эффективность и производительность в каком-то виде деятельности, то это говорит о том, что деятельность автоматизирована и далее не представляет ценности для трат мыслительных ресурсов человека.



# Спасибо за внимание!

Контактная информация:

Карауш Александр Сергеевич

e-mail: [a@karaush.ru](mailto:a@karaush.ru) / [ask@gpntb.ru](mailto:ask@gpntb.ru)

telegram: [akaraush](https://www.instagram.com/akaraush)