



Data-as-a-Service (DaaS) to create analytical tools to drive effective R&D and deliver monetizable innovation

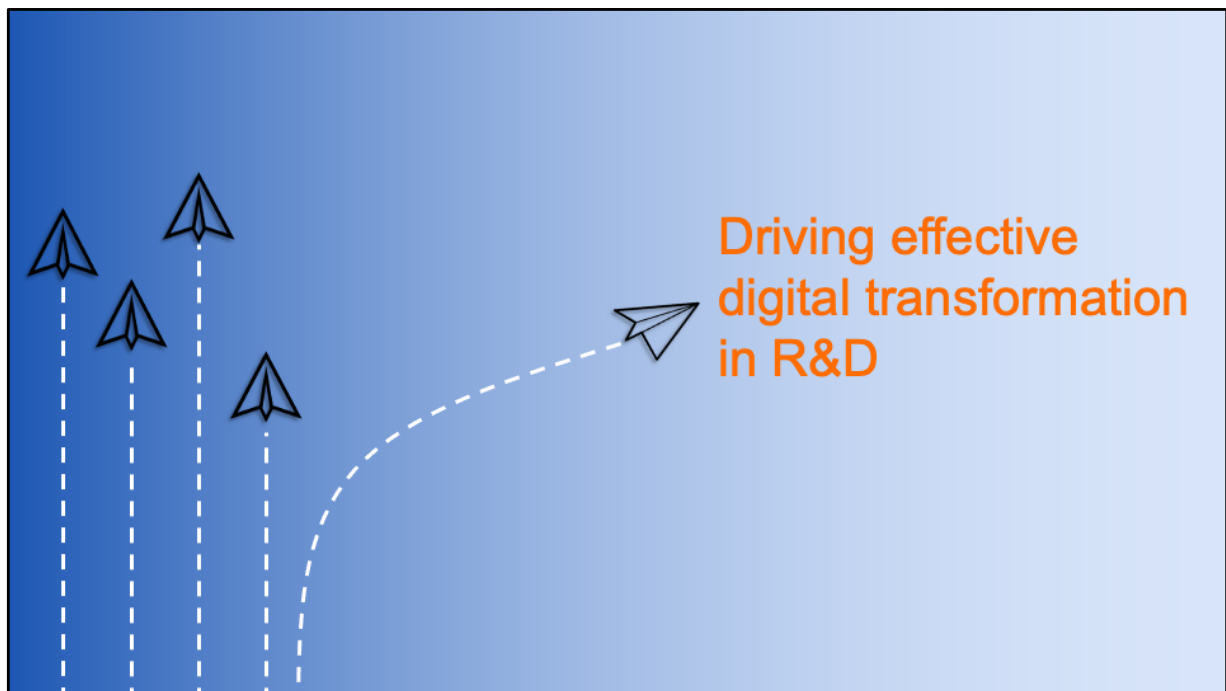
November 20, 2020

Jabe Wilson PhD MBA

Global Commercial Director, Data & Analytics, Corporate R&D Markets



Презентация на тему: “Данные-как-услуга (DaaS)”: сервис для создания аналитических инструментов поддержки эффективных НИОКР и внедрения коммерчески успешных инноваций



- Производительность традиционных лабораторных исследований и разработок в последние 20 лет снижается
- В отраслях, интенсивно использующих НИОКР, происходит целенаправленное изменение способов выполнения НИОКР в сторону концентрации ресурсов на исследованиях in-silico, основанных на данных
- В результате этой цифровой трансформации ожидается улучшение ключевых показателей бизнеса в индустрии (снижение затрат, монетизируемые инновации, ускоренный выход на рынок)
- Эффективная цифровая трансформация по-прежнему является проблемой для многих организаций
- Цифровая трансформация - это технология + данные
- Эффективная стратегия работы с данными основана на использовании множества типов данных, как внутренних, так и внешних,
- Для реализации потенциала цифровизации необходимо партнерство организаций, каждая из которых может извлечь свою выгоду из данных
- Elsevier предлагает партнерство, которое поможет компаниям найти необходимые данные и использовать их в процессе цифровой трансформации

Elsevier serves the research and digital transformation needs of a diverse set of research and engineering intensive industries

Elsevier serves thousands of corporate customers globally across multiple industries, including:

 Pharmaceuticals & Biotechnology	 Chemicals & Materials	 Oil & Gas
 Electronics & Semiconductors	 Aerospace & Defence	 Engineering & Construction



Elsevier обслуживает тысячи корпоративных клиентов по всему миру в самых разных отраслях:

- Фарма и биотехнологии
- Химия и материалы
- Нефть и газ
- Электроника и компоненты
- Аэрокосмическая отрасль
- Строительство и инжиниринг



Who We Are
Global information analytics business
specializing in science and technology

Why We Do It
To help solve your
challenges, for the benefit
of the scientific community

What We Do
Help scientists and researchers
improve performance and
advance science

A Unique Combination
Combine content with technology, supported
by operational efficiency, to turn information
into actionable knowledge.



- Наша компания занимается анализом информации в области науки, технологии и медицины
- Мы работаем над проблемами наших клиентов путем сочетания интеллектуальных технологий с контентом и производственными ноу-хау

Today we focus on 3 examples

1. Improving energy efficiency in steel production
2. Chemical synthesis with reduced hazardous side products
3. Drug repurposing for disease treatment in healthcare



Сегодня мы рассмотрим 3 примера использования нашего подхода:

- Повышение энергоэффективности при производстве стали
- Химический синтез с пониженным содержанием вредных побочных продуктов
- Перепрофилирование лекарств для лечения заболеваний

Digital transformation for more effective and efficient delivery of R&D outcomes



Launch new products or extend product coverage

Identify the right growth opportunities through shared expertise and use modelling data to find opportunities for patent expansion.



Reduce R&D time to market

Make quicker decisions about which therapies or products to advance through the research pipeline. Drive more effective R&D at each stage with data-driven decisions.



Increase R&D cost efficiencies

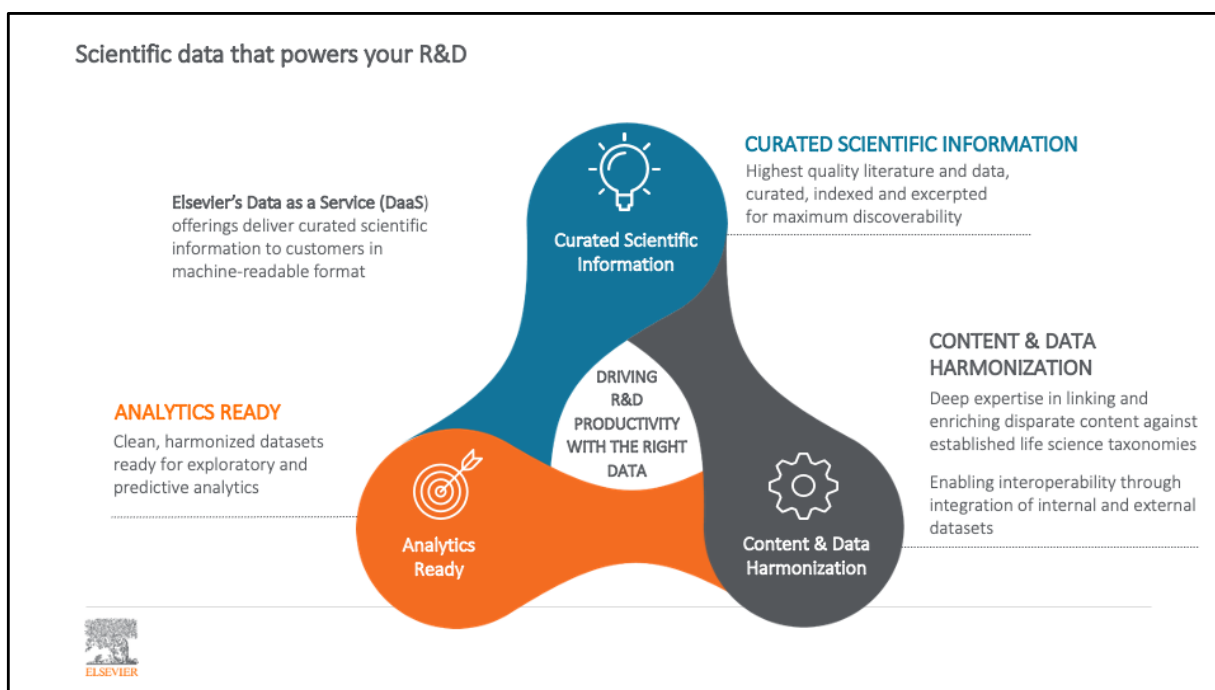
Perform in silico research before real-world experimentation. Model many candidates using predictive analytics before selecting the few to model in the lab to reduce unnecessary experimentation duplication.



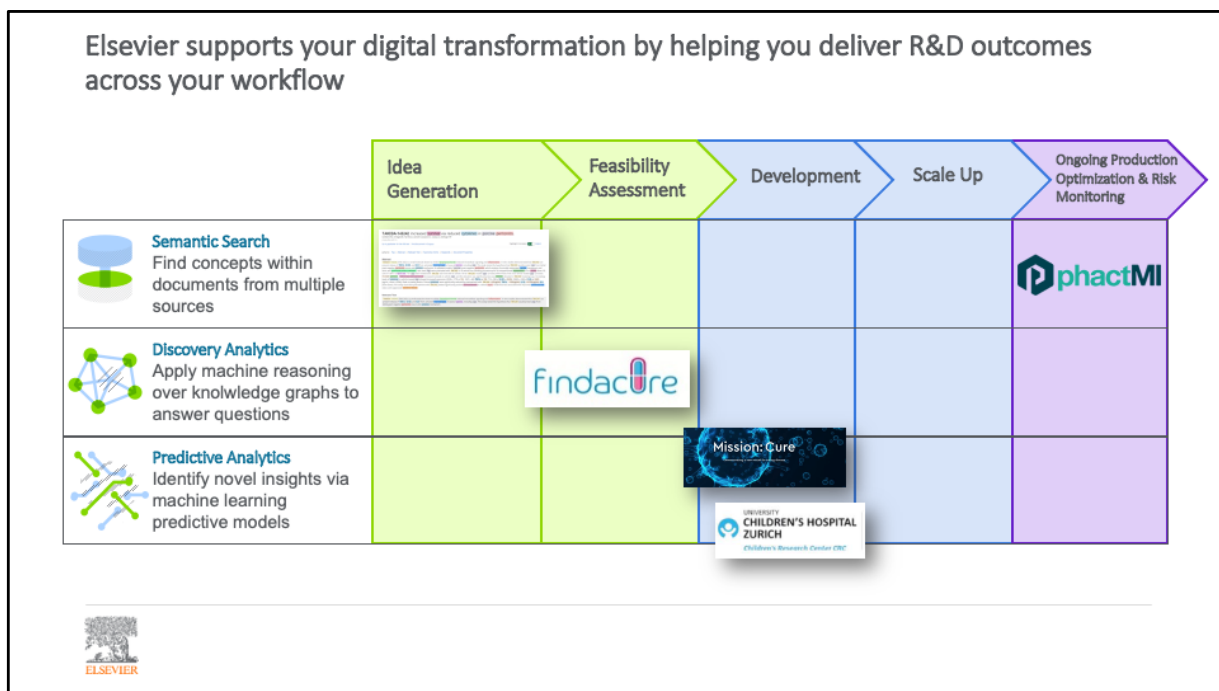
Inspiration: Slide 11, <https://www.pwc.es/publicaciones/pharma/assets/pharma-2020-de-la-vision-a-la-decision.pdf>



- Индустрия активно использует ИИ для повышения эффективности и производительности
- Ситуация с COVID только ускорила процесс, «нажали на педаль газа»
- Компании, инвестирующие в цифровую трансформацию, получают выгоду 3-мя способами:
 - Генерация новых идей
 - Сокращение времени на исследования и разработку новых продуктов
 - Повышение рентабельности инвестиций в НИОКР
- Цифровая трансформация сложна - много переходных проблем, связанных с людьми, технологиями и данными.
- Мы видим несколько проблем, которые необходимо решить:
 - Объем публикуемых данных увеличивается с каждым годом, нет возможности отслеживания даже самым подготовленным исследователям
 - Большая неоднородность научных областей и соответствующих типов данных, что приводит к сложностям, когда группы НИОКР пытаются согласовать и установить значимые связи между данными
 - Данные часто хранятся разрозненно, особенно в организациях с долгой историей традиционных экспериментов.
 - Дублирование усилий: исследователи часто повторяют работы коллег в своей компании или опубликованные исследования, доступные извне.

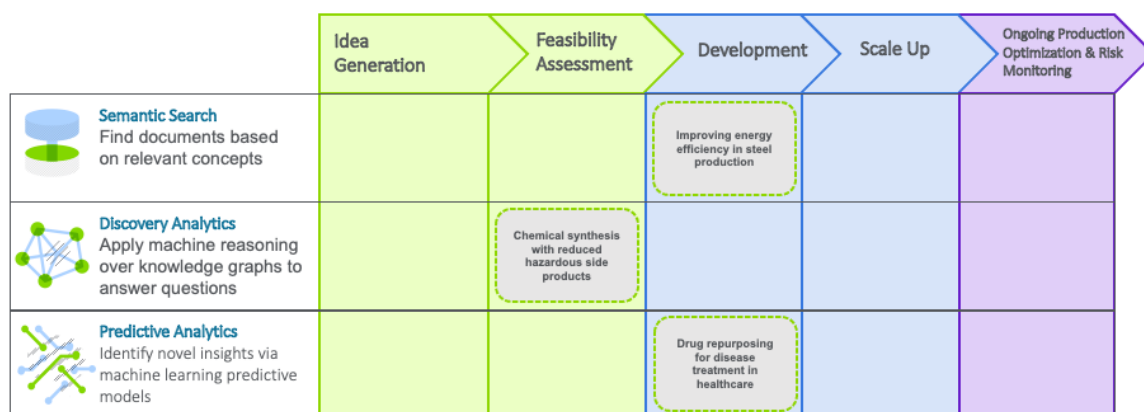


- Elsevier помогает компаниям повысить продуктивность их исследований и разработок за счет предоставления доступа к нужным данным в нужном формате в нужное время
- Elsevier является одним из самых надежных поставщиков структурированных и неструктурированных научных данных, что является ключевым фактором для успешной цифровой трансформации
- Мы ищем новые способы сотрудничества, чтобы улучшить результаты для наших клиентов.... одним из них является наш новый сервис «Данные как услуга» (Data-as-a-service)
- Data-as-a-service означает, что вы сможете использовать данные за пределами исходной базы (контейнера)
- Мы предоставляем не только данные, но и опыт решения конкретных задач, получая неструктурированные и структурированные данные из широкого спектра источников и дисциплин для создания надежных, структурированных (FAIR) источников данных.
- Эти источники данных используются для прикладной аналитики и прогнозов - определение новых путей лечения, создание платформ для выявления трендов и лидеров мнений в материаловедении и многие другие приложения



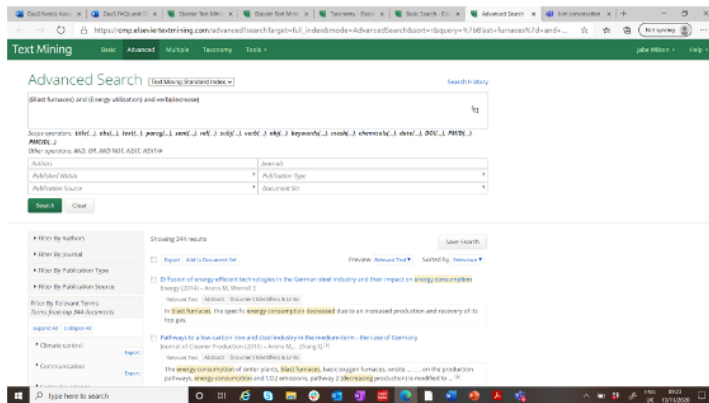
- Мы активно работаем с компаниями, ставящими на R&D в качестве основного элемента стратегии развития
- Акцент на достижение реальных результатов в цепочке создания стоимости исследований (в любых отраслях), как пример:
 - предоставление специалистам здравоохранения надежной и точной медицинской информации для точных рецептов (например, в рамках нашего партнерства с phactMI)
 - выявление потенциальных возможностей перепрофилирования лекарств
 - прогнозирование возможностей эффективного проектирования производственного процесса путем моделирования взаимодействия материалов
- Хорошие данные помогают достичь ощутимых результатов, а достоверные данные также помогают убедить группы НИОКР использовать инструменты, разработанные группами науки о данных и in-silico.

Elsevier supports your digital transformation by helping you deliver R&D outcomes across your workflow



- Мы активно работаем с компаниями, ставящими на R&D в качестве основного элемента стратегии развития
- Акцент на достижение реальных результатов в цепочке создания стоимости исследований (в любых отраслях), как пример:
 - предоставление специалистам здравоохранения надежной и точной медицинской информации для точных рецептов (например, в рамках нашего партнерства с phactMI)
 - выявление потенциальных возможностей перепрофилирования лекарств
 - прогнозирование возможностей эффективного проектирования производственного процесса путем моделирования взаимодействия материалов
- Хорошие данные помогают достичь ощутимых результатов, а достоверные данные также помогают убедить группы НИОКР использовать инструменты, разработанные группами науки о данных и in-silico.

Use case: Energy reduction in blast furnaces



[Live link in Elsevier Text Mining](#)

- Здесь мы показываем, как семантический поиск через ScienceDirect может предоставить конкретные статьи, в которых обсуждается снижение потребления энергии в доменных печах

Use case: Energy reduction in blast furnaces

The screenshot displays the Elsevier Text Mining web application. The top navigation bar includes 'Text Mining', 'Basic', 'Advanced', 'Multiple', 'Tutorials', and 'Tools'. The main content area is titled 'Advanced Search' and shows search results for the query '(blast furnace) and (energy utilization and energy consumption)'. The search results list the article 'Diffusion of energy efficient technologies in the German steel industry and their impact on energy consumption' by Anais M. Worrell, published in Energy (2014). The article abstract is visible, discussing the role of technological change and the diffusion of energy efficient technologies in the German steel industry. The abstract mentions that the authors selected six key energy efficient technologies and collected data to show their diffusion since their introduction in Germany. The technologies are: blast furnace gas (BFG) recovery, hot blast gas (HBG) recovery, hot blast gas (HBG) recovery, hot blast gas (HBG) recovery, hot blast gas (HBG) recovery, and hot blast gas (HBG) recovery. The abstract also mentions that the authors found that the diffusion of these technologies was significantly higher in the German steel industry compared to other countries. The abstract concludes that the diffusion of these technologies has led to a significant reduction in energy consumption in the German steel industry.

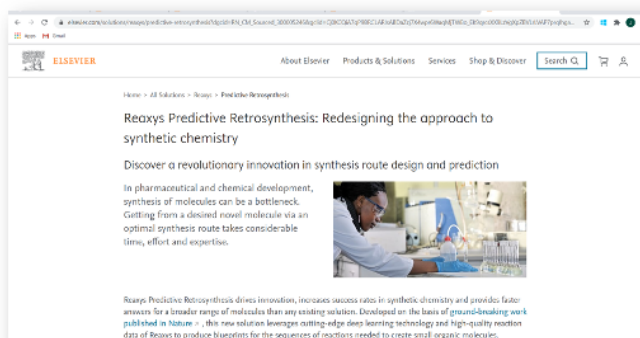


[Live link in Elsevier Text Mining](#)

На этом слайде показано одно из резюме статьи с выделенными соответствующими терминами.

Retrosynthesis with deep neural networks and symbolic AI

<https://www.elsevier.com/solutions/reaxys/predictive-retrosynthesis>



Marwin H. S. Segler^{1,2}, Mike Preuss³ & Mark P. Waller⁴

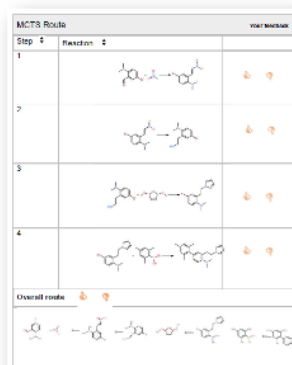
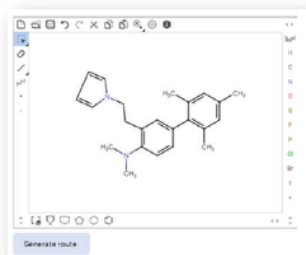
<https://doi.org/10.1038/nature25978>

<https://www.elsevier.com/solutions/reaxys/predictive-retrosynthesis>

На этом слайде обсуждается применение революционного алгоритма искусственного интеллекта для прогнозирования маршрутов химического синтеза, обученного на данных Elsevier Reaxys

Retrosynthesis with deep neural networks and symbolic AI

Draw a molecule in Marvin JS



The first part shows reaction formula of each step.
The **"Overall route"** part presents a general sketch.



Marwin H. S. Segler^{1,2}, Mike Preuss³ & Mark P. Waller⁴

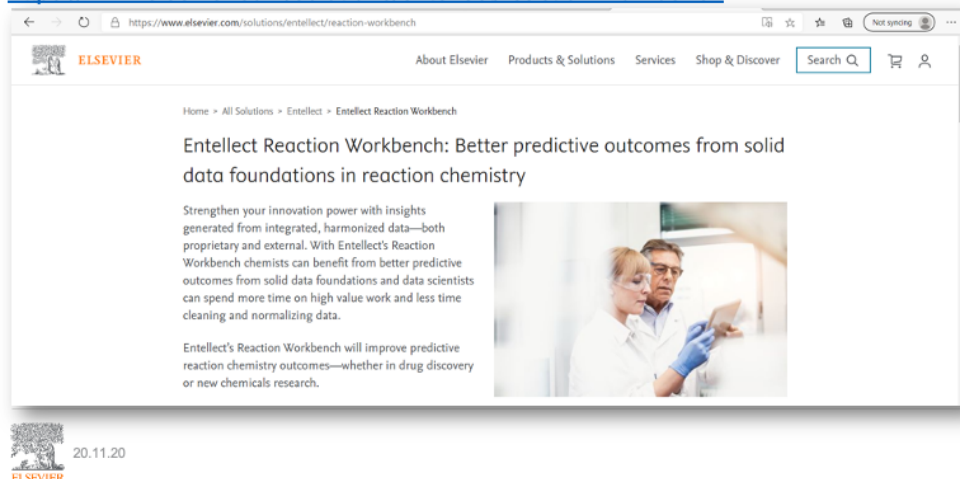
<https://doi.org/10.1038/nature25978>

<https://www.elsevier.com/solutions/reaxys/predictive-retrosynthesis>

Вы можете увидеть, как это выглядит на экране: вы рисуете молекулу, которую хотите создать, и алгоритм планирует маршрут.

Retrosynthesis: green route and sustainable products

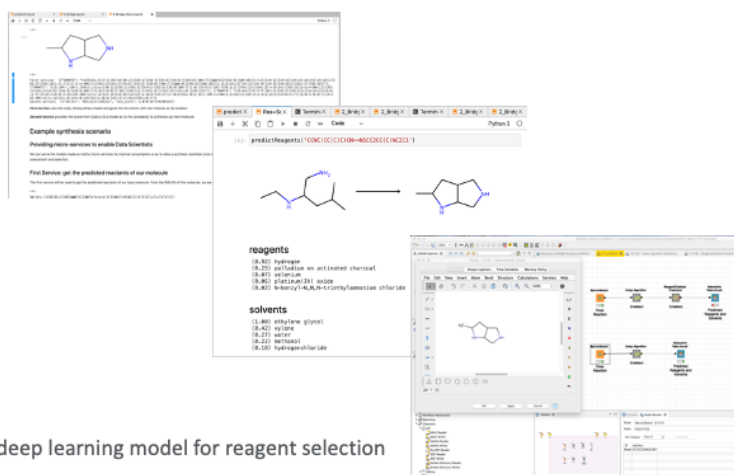
<https://www.elsevier.com/solutions/entellect/reaction-workbench>



Elsevier также предоставляет среду искусственного интеллекта, в которой вы можете создавать, редактировать и развертывать прогнозные модели, которые можно настроить, чтобы избежать опасных побочных продуктов.

Retrosynthesis: green route and sustainable products

<https://www.elsevier.com/solutions/entellect/reaction-workbench>



Вы можете увидеть, как эти модели могут быть развернуты в популярном инструменте информатики, таком как Knime.

AI predictions for Drug Discovery in Healthcare



- **Step 1).** Data preparation: Getting your data together



- **Step 2).** Building a knowledge graph: Using different lenses to see the data clearly



- **Step 3).** AI ready data: Creating slices of data to feed Machine Learning



- **Step 4).** Creating an AI tool: Running Machine Learning to make a trained Artificial Neural Network



- **Step 5).** Responsible AI: Putting reliable predictions in the hands of the experts



- Ключевые шаги в создании инструмента прогнозирования для любого варианта использования НИОКР - на примере перепрофилирования лекарств, который недавно получил престижную награду за инновации BioIT World (Джейб Уилсон недавно представил ее члену Европейского союза. Парламент сосредоточился на этическом использовании ИИ)

- Шаг первый: подготовьте свои данные

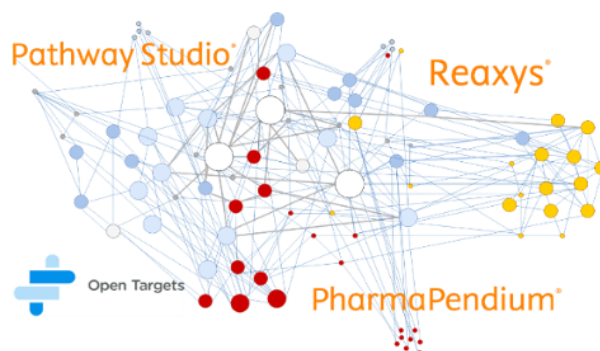
- Шаг второй: связывайте факты с помощью таксономий

- Шаг третий: преобразование данных в векторы (серии единиц и нулей, на которых можно обучать машинное обучение)

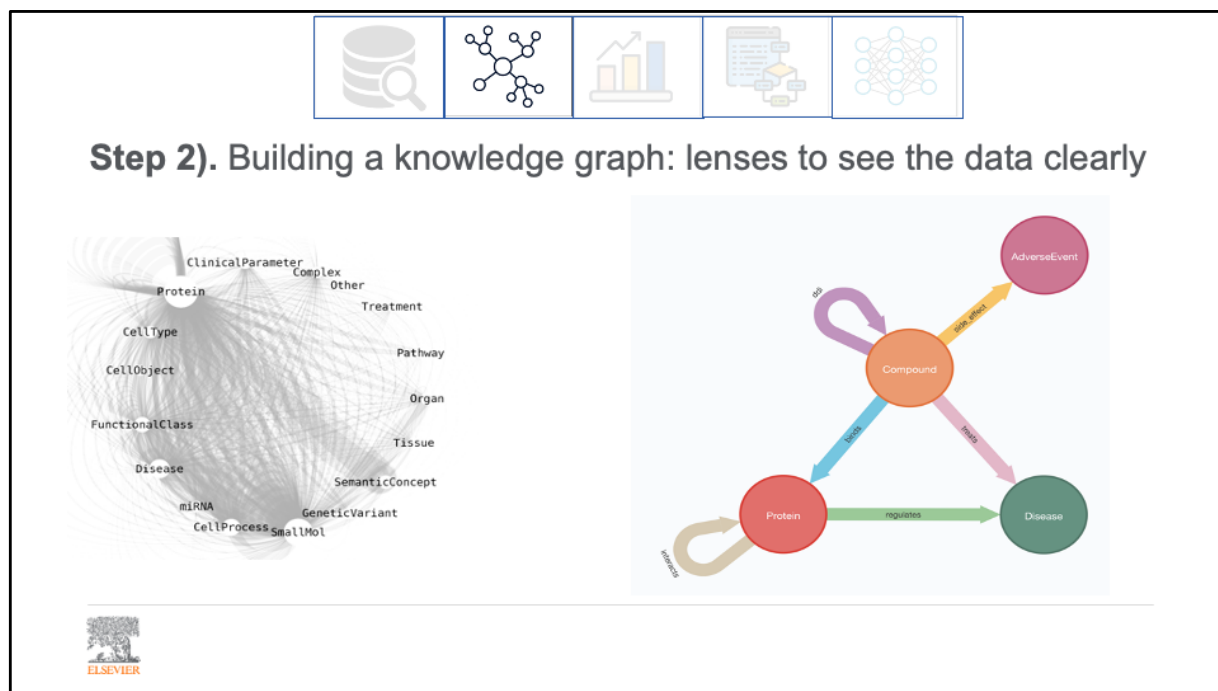
- Шаг четвертый: запустите алгоритм машинного обучения и создайте прогнозирующую нейронную сеть.



Step 1). Data preparation: Getting your data together



В качестве примера - это проект, который взял данные из Elsevier и сторонних источников и интегрировал их в базу данных графа знаний.



В результате получился большой сложный график, который затем был сокращен путем создания упрощенной модели данных, ориентированной на данные, необходимые для перепрофилирования лекарств.

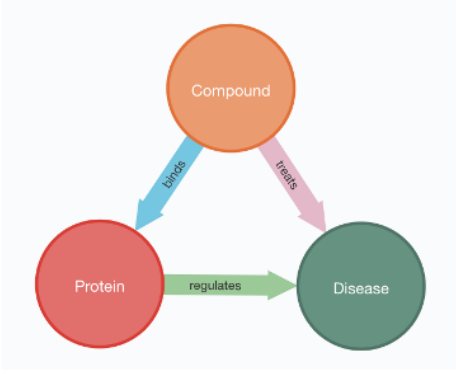


Step 3). AI ready data: Creating slices of data to feed Machine Learning

Strategy 1: Can we predict new drivers of the disease?

Strategy 2: Can we predict off-targets interactions of a drug that are genetic drivers of the disease?

Strategy 3: Can we predict if a drug treats a disease via any drug target and any mechanism?



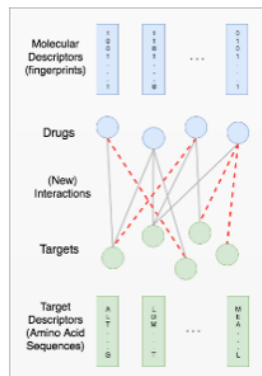
```
graph TD; Compound((Compound)) -- binds --> Protein((Protein)); Protein -- regulates --> Disease((Disease)); Compound -- treats --> Disease;
```



Для перепрофилирования лекарств часто используются три стратегии - этот проект сосредоточен на второй стратегии - прогнозировании генов экспрессируемых белков, участвующих в заболевании, с которыми лекарственные соединения могут связываться и останавливать болезнь.



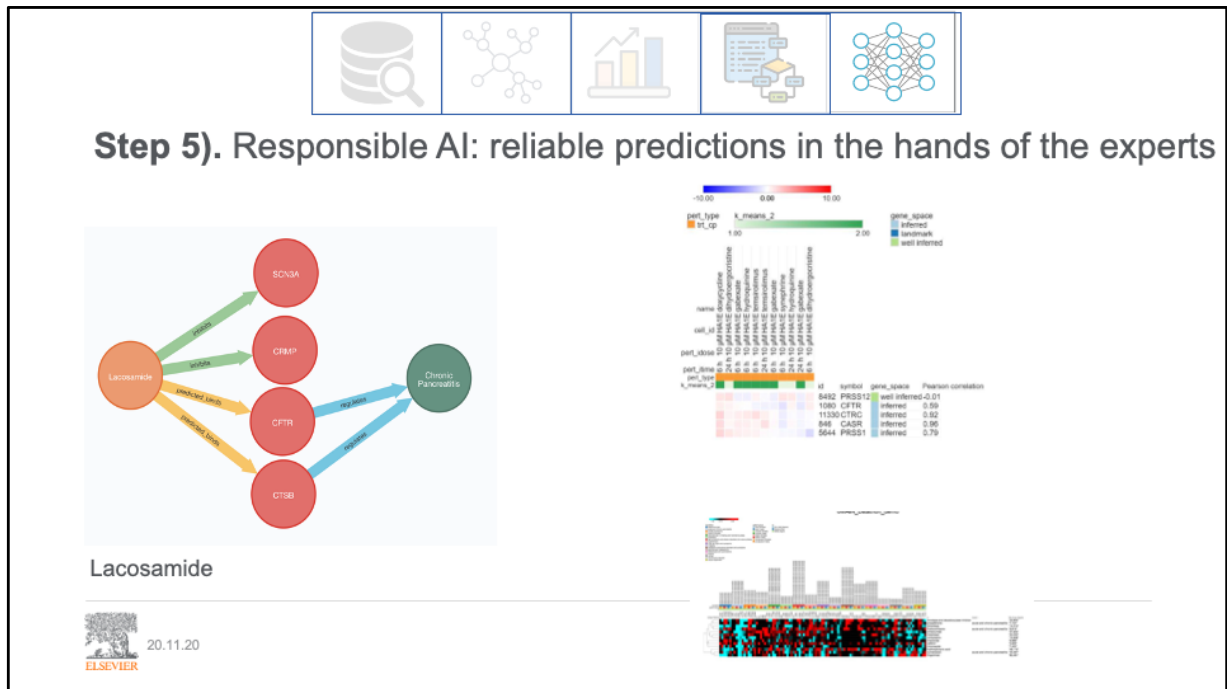
Step 4). Creating an AI tool: Running ML to train a Neural Network



- Vectors representations (proteins and substances)
- More than one ML model approach - dimensionality reduction - large data (120k interactions, 44k substances, 200 gene-encoded proteins)



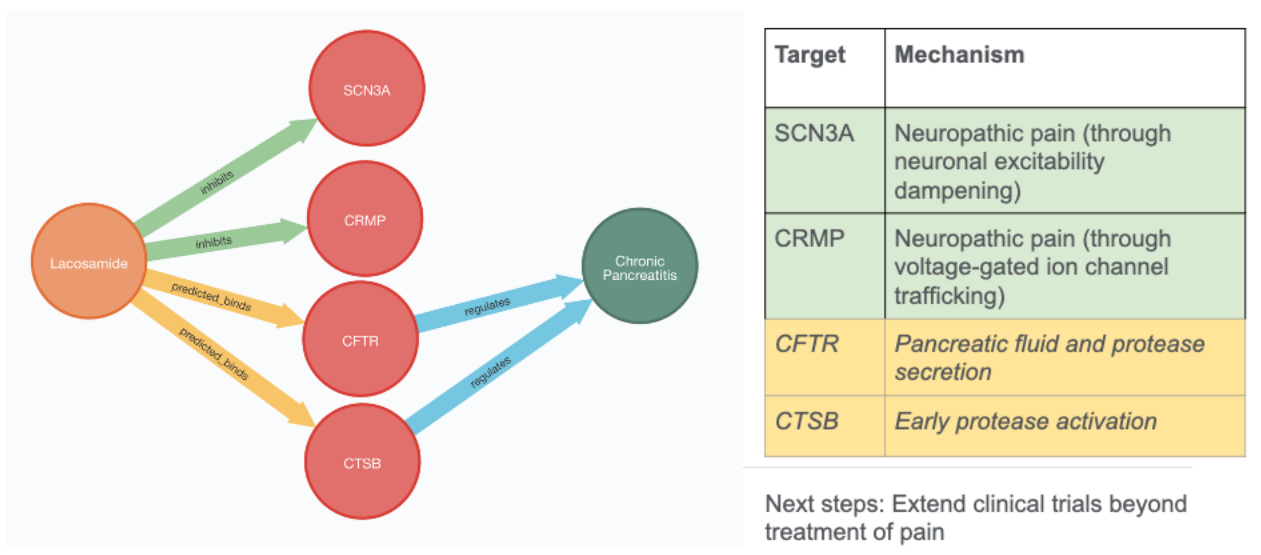
Следующим шагом было необходимо запустить машинное обучение на основе данных, готовых к ИИ, для прогнозирования новых связываний лекарств с белками (то есть новых методов лечения).



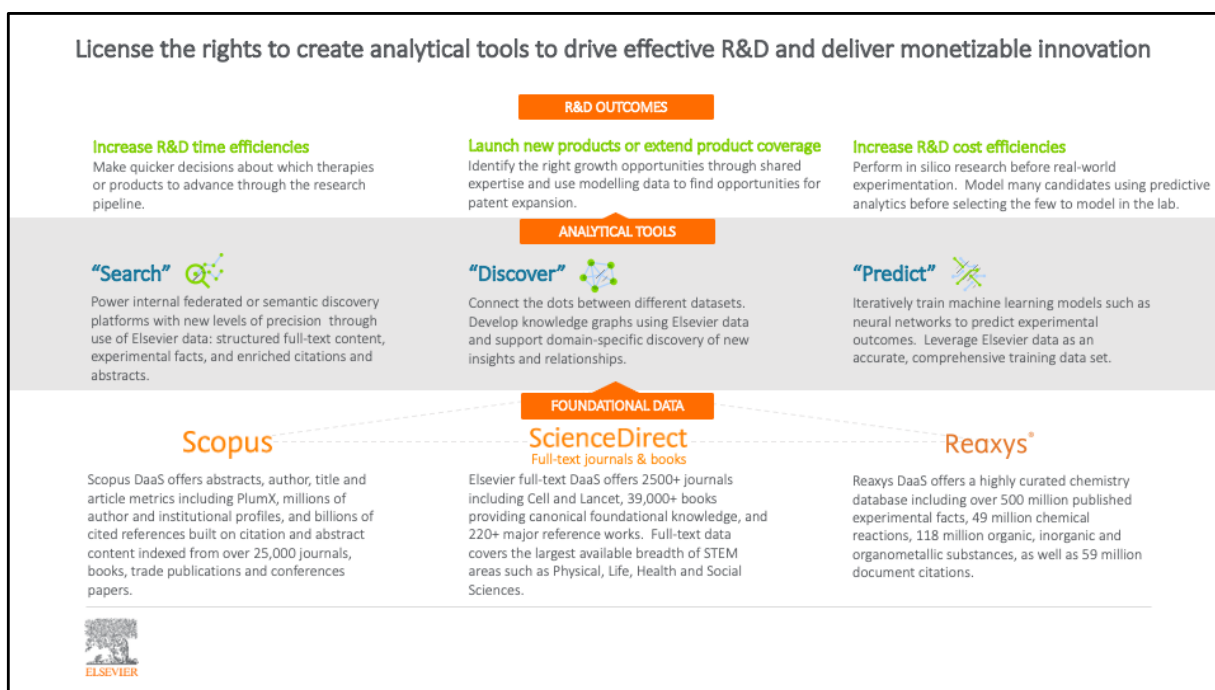
Рекомендованный препарат Лакозамид был передан медицинским экспертам, которые провели дополнительный анализ для оценки рекомендованного препарата.

Lacosamide for clinical trial

Strategy 2: Can we predict off-targets interactions of a drug that are genetic drivers of CP?



- Имея как генетические мишени, так и обученные модели, мы провели виртуальный скрининг лекарств, чтобы предсказать вероятность того, что каждое лекарство, одобренное FDA или EMA, связывается с генетическим драйвером CP дикого типа
- Эти прогнозы послужили катализатором для обсуждений с малыми и средними предприятиями и дальнейшего анализа, включая омики и сопутствующие неблагоприятные события
- Мы составили список и обнаружили, что лакозамид является наиболее интересным
- Вычислительные методы предсказали, что лакозамид связывается как с CFTR, так и с CTSB
- Доклинически доказано, что ингибирование CTSB предотвращает раннюю активацию протеазы.
- В настоящее время нет одобренных ингибиторов CTSB, в большинстве исследований используется делеция гена, поэтому сравнительный анализ с этим труднее проводить.
- Помимо прогнозируемых нецелевых целей лакозамида, которые необходимо будет проверить экспериментально, известные целевые объекты могут оказаться терапевтически эффективными против ХП.
- Известный компонент боли в животе при ХП - невропатический. Исследования показали, что лакозамид может снизить чувствительность нервных волокон и уменьшить восприятие боли пациентами.
- Таким образом, он был предложен в качестве лечения нескольких состояний с компонентом невропатической боли, включая клинические испытания фазы 2, которые в настоящее время набираются для лечения невропатии.



- Сервис DaaS обеспечивает огромный объем надежного высококачественного машиночитаемого контента и данных для всей команды НИОКР, стимулируя взаимодействие между лабораторными исследованиями, in-silico и наукой о данных
- Мы можем сотрудничать с вами или напрямую лицензировать вам права на данные
- Вы можете создавать свои собственные аналитические инструменты (и владеть IP) на основе наших данных
- Scopus включает в себя аннотации и индексирующие данные по предметам STEM, ScienceDirect охватывает полнотекстовые статьи из ведущих журналов и книг во всех областях науки, Reaxys содержит беспрецедентный объем тщательно отобранных химических реакций, структур и ссылок на документы
- Эти данные могут применяться следующим образом:
 - **Поиск:** внеплатформенные данные Elsevier могут быть включены в ваши внутренние федеративные или семантические платформы поиска
 - **Отслеживание:** специалисты по данным и информатике смогут связать разные наборы данных и разрабатывать графы знаний (набор взаимосвязанных описаний сущностей - в контексте)
 - **Прогнозирование:** вы можете итеративно обучать и строить модели машинного обучения, такие как нейронные сети, для прогнозирования результатов экспериментов.

What it means to you? Call to Action

- License the rights to data to create analytical tools
- Professional Services
- SciBite AI licenses

Discuss your needs further? Please contact: jabe.wilson@elsevier.com



- Elsevier может поддержать вашу цифровую трансформацию в исследованиях и разработках путем разработки таких аналитических инструментов
- Мы можем сделать это одним из трех способов:
 - Мы можем лицензировать ваши данные и права создавать аналитические инструменты на основе данных
 - Наша команда профессиональных услуг может работать для вас над проектами по созданию таких инструментов
 - Мы также можем лицензировать вам программные инструменты SciBite для достижения этих результатов.
- Приглашаем к конкретному обсуждению Ваших задач



Thank you

