

ОТЕЧЕСТВЕННАЯ ПЛАТФОРМА ИИ ДЛЯ ОРГАНИЧЕСКОЙ И МЕДИЦИНСКОЙ ХИМИИ

фцприи.рф

Проблемы:

1 Из-за высоких расходов на эксперименты, стоимость разработки новых лекарственных средств непрерывно увеличивается

2общая частота неудач при разработке лекарств составляет **более 96%**, включая 90% неудач при **клинической разработке***...

Следствия для потребителя:




Уход от реальных **инноваций к производным** уже существующих соединений



Возрастание **себестоимости** новых лекарств

**Разработка нового лекарства —
крайне долгий и дорогой процесс**

от 0 лет \approx 1 млрд.  **до 10 лет**

1 Токсичность

2 Недостаточная эффективность
молекулы-кандидата

30%

*FDA утверждает

*новых кандидатов
в лекарственные препараты не проходят 1 стадию
клинических испытаний (небезопасны)*

ОТБОР МОЛЕКУЛ-КАНДИДАТОВ (ВОРОНКА КАНДИДАТОВ)



Огромный поток кандидатов на входе



Отбор



Идеальный кандидат

ЭКСПЕРИМЕНТАЛЬНЫЕ МЕТОДЫ НЕ СПРАВЛЯЮТСЯ С БОЛЬШИМ КОЛИЧЕСТВОМ НОВЫХ МОЛЕКУЛ

1 Измерение полного набора физико-химических параметров для **одного соединения требует нескольких месяцев лабораторной работы и \$50-100K за соединение!**

2 Компании из фармацевтического, агрохимического, пищевого и прочих секторов **тратят на эти исследования миллиарды долларов в год!**

Разработанный нами подход **автоматического** предсказания свойств химических соединений позволяет сократить на **несколько порядков** временные и денежные затраты.

Physical

| | | |
|-------------------------|---------------------------|-----|
| Water Solubility | -1.0333232 LogS | EXP |
| Water Solubility (QSAR) | -1.511 ± 0.095 Log(mol/L) | |
| Vapor Pressure | -4.2 log(mmHg) | |
| Vapor Pressure (QSAR) | -19.42 ± 9.67 Torr | |
| Boiling point | 296 Celsius | |
| Boiling point (QSAR) | 322.7 ± 16.7 °C | |
| Flash point | 188.4 Celsius | EXP |
| Density | 1.249 g/cm3 | EXP |
| Density (QSAR) | 1.298 ± 0.01 g/cm3 | |
| Density (normal) (QSAR) | 1.308 ± 0.017 g/cm3 | |
| Viscosity | 0.9 log10(viscosity cP) | |
| Melting point | 170 Celsius | EXP |
| Melting point (QSAR) | 170.6 ± 3.6 °C | |
| LogP octanol-water | 0.46 logP | EXP |
| Soluble in DMSO | Soluble | EXP |
| DMSO Solubility (QSAR) | 1 0 (Soluble) | |
| Retention time | 430 s | |
| Retention time (QSAR) | 523.5 ± 54.3 s | |
| Refractive Index (QSAR) | 1.565 ± 0.006 | |

Химическое пространство огромно. Его объем оценен в 10⁶⁰ малых молекул. Искать необходимые молекулы в таком огромном массиве вручную - довольно затруднительно. Экспертам необходимы современные методы навигации в химическом пространстве. Искусственный интеллект может обеспечить такую возможность.



Лекарственные
средства



Топливо и
масла



Пластмассы и
композиты

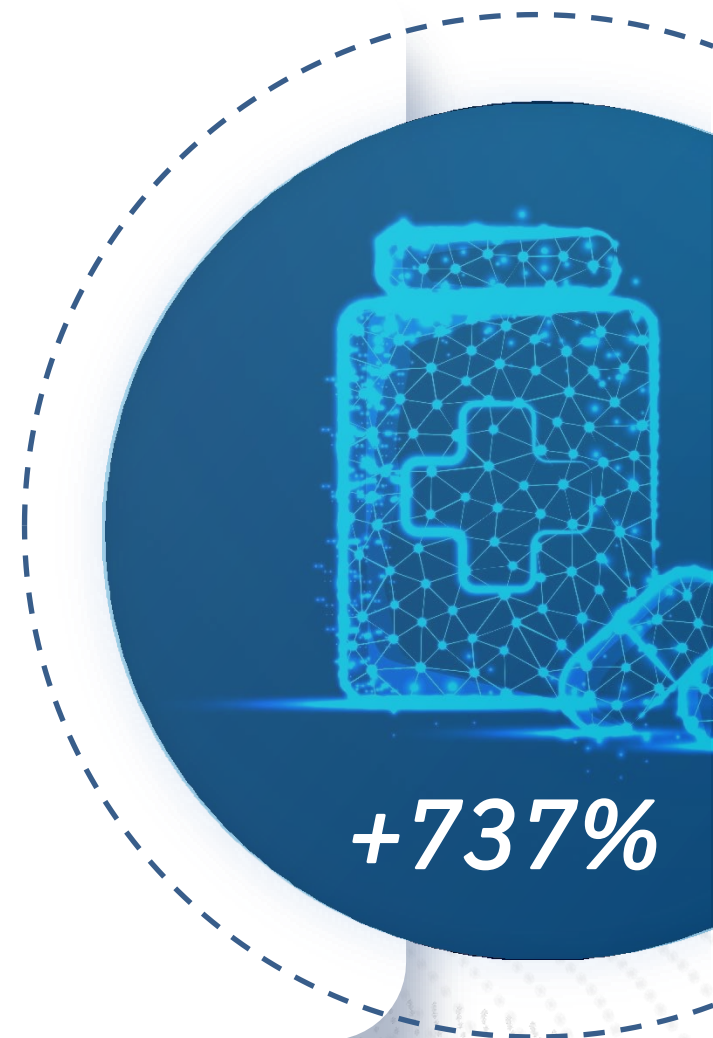
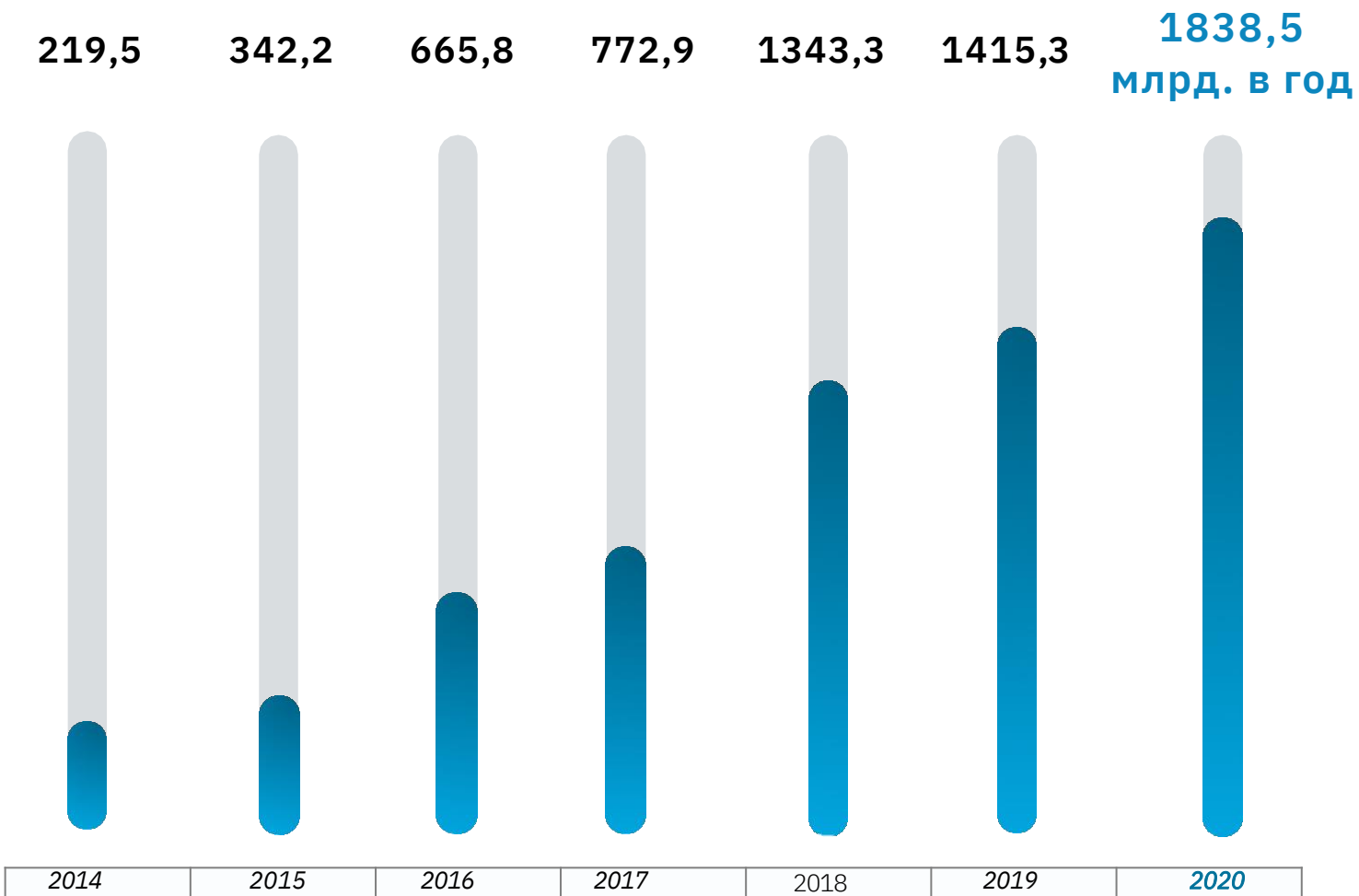


Жидкие кристаллы и
OLED-технологии

4 **\$** **МЛРД.** **Объем мирового
рынка
хемоинформатики
в 2016 году***

- Эпидемии приведут к стремительному росту вложений в сферу разработки лекарств;
- Искусственный интеллект показывает крайне высокую эффективность в задачах хемоинформатики (возможно решить "нерешаемые" ранее задачи).

РОСТ ИНВЕСТИЦИЙ В ИИ ЗА ПОСЛЕДНИЕ 7 ЛЕТ В ОБЛАСТИ РАЗРАБОТКИ ЛЕКАРСТВ



Программный комплекс хемоинформатики на базе искусственного интеллекта для решения задач органической и медицинской химии.

МОДУЛЬНАЯ ПЛАТФОРМА ВКЛЮЧАЕТ В СЕБЯ:

Инструменты для визуальной навигации по
химическому пространству

Анализ кластеров биоактивных соединений

01

Модуль прогнозирования исхода реакций

Анализ синтеза искомой молекулы

02

Оптическое распознавание
молекулярных структур из PDF

Извлечение химических данных из
различных неструктурированных
печатных источников

03

ПРОГРАММНЫЙ КОМПЛЕКС ВКЛЮЧАЕТ В СЕБЯ:

Базу данных

Поиск необходимых молекул и быстрый доступ к их
свойствам. Основная база содержит 96 млн. молекул;

01

Генератор новых структур

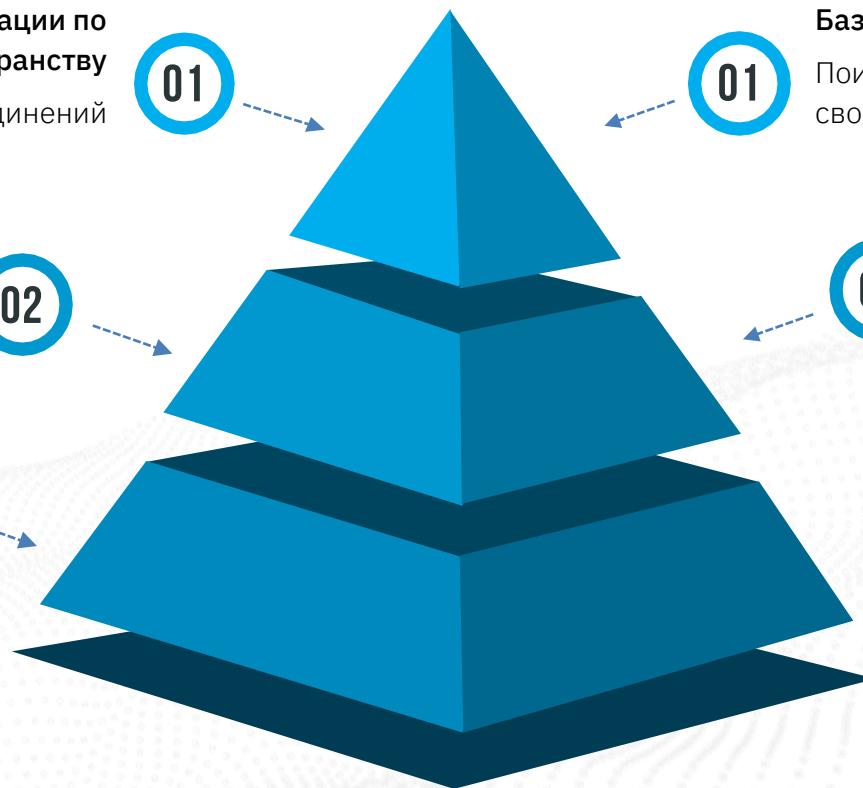
Конструктор соединений с заранее заданными
пользователем набором свойств (низкая токсичность,
низкая синтетическая сложность, высокая
биологическая активность);

02

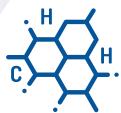
Набор прогностических моделей на основе
глубоких нейронных сетей

Для расчета физико-химических,
токсикологических, биологических и
экологических свойств органических
соединений;

03



Разработка лекарственных препаратов в сегментах оригинальных препаратов и перепрофилирования лекарственных средств на основе малых молекул, включая уникальную функцию предсказания токсичности и биологической активности, а также сложности синтеза. Наибольшая ценность платформы возникает в случае эпидемий, когда требуется разработка новых препаратов в кратчайшие сроки.



Химическая промышленность

в части разработки новых соединений и материалов для нефтехимии, пестицидов, гербицидов и производства пищевых добавок и т.д., с комплексной оценкой безопасности соединений и промежуточных реагентов.



Косметическая промышленность

в части разработки новых безопасных композиций.



Защита интеллектуальной собственности и патентный поиск (для патентования новых химических соединений).



Разработка лекарственных препаратов

(в сегментах оригинальных препаратов и перепрофилирования лекарственных средств на основе малых молекул), включая уникальную функцию предсказания токсичности и биологической активности, а также сложности синтеза. Наибольшая ценность платформы возникает в случае эпидемий, когда требуется разработка новых препаратов в кратчайшие сроки.



Регуляторная деятельность в части контроля и **проверки безопасности** (токсичности) химических веществ.

ПЛАТФОРМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ИЗУЧЕНИЯ ХИМИЧЕСКОГО ПРОСТРАНСТВА,
ПРОГНОЗИРОВАНИЯ СВОЙСТВ ОРГАНИЧЕСКИХ СОЕДИНЕНИЙ И РЕШЕНИЯ ЗАДАЧ МЕДИЦИНСКОЙ ХИМИИ.

МОЛЕКУЛЫ

Ищите необходимые соединения и получайте быстрый доступ к информации об их свойствах. Основная база содержит более 96 миллионов молекул;

ОРГАНИЧЕСКИЕ
РЕАКЦИИ

Прогнозируйте возможные продукты химических реакций. Осуществляйте поиск реакций для синтеза искомой молекулы;

МОЛЕКУЛЯРНАЯ
КАРТА

Исследуйте химическое пространство, используя наши модели глубокого обучения. Анализ кластеров биоактивных соединений. Генерация новых соединений с заданными свойствами;

ИНДИВИДУАЛЬНЫЙ
РЕЖИМ

Используйте блок прогнозирования свойств для органических молекул, которых нет в нашей базе: физико-химические, биологические свойства, токсичность, канцерогенность, сложность синтеза, прогноз IUPAC имен;

PDF2SMILES

Используйте технологии оптического распознавания молекулярных структур и структур Маркуша из PDF.

ОБЕСПЕЧЕНИЕ ТЕХНОЛОГИЧЕСКОГО СУВЕРЕНИТЕТА ТРЕБУЕТ УСТРАНЕНИЯ ПРОБЕЛА В ОБЛАСТИ ХРАНЕНИЯ, ПРЕДОСТАВЛЕНИЯ И АНАЛИЗА ХИМИЧЕСКОЙ ИНФОРМАЦИИ

- 01** Ограничения доступа к подобной инфраструктуре для российских специалистов для российских ученых при разработке лекарственных препаратов, новых веществ, материалов и исследования их свойств **ведет к критическому отставанию и практической остановке многих исследований.**
- 02** В РФ отсутствует единая национальная база данных, включающая набор стандартов в области хранения, предоставления и анализа химической информации, аналогичная **PubChem (US) или ChemSpider (UK)**, что ведет к потере суверенитета в области химических исследований.
- 03** Уход из России ряда провайдеров зарубежного специализированного ПО и баз данных в области химии и фармакологии, критическая зависимость российского научного сообщества от этих сервисов.
- 04** Экспериментальный поиск соединений, обладающих заданными свойствами, является крайне дорогостоящим, цена синтеза и профилирования физико-химических свойств только одной молекулы может достигать до **\$100 000**. Необходимость исследования токсикологических свойств увеличивает эту сумму **в несколько раз.**
- 05** На глобальном рынке разработка нового медицинского препарата занимает в среднем **10-12 лет** и требует инвестиций от **\$1 млрд**. Стоимость разработки новых лекарственных средств непрерывно увеличивается, при этом процент неудач в разработке достигает **96%.**
- 06** Острая необходимость вследствие геополитической ситуации, в укреплении научного и технологического суверенитетов и в создании собственной инфраструктуры для их поддержки.

| Объем рынка | Химическая промышленность (сегмент новых соединений и материалов) | Разработка фарм. препаратов (сегмент оригинальных препаратов и дженериков на основе малых молекул) | Косметическая промышленность | Университеты и НИИ (химия, фармакология, токсикология) | Патентные организации и палаты патентных поверенных | Технопарки, инкубаторы, стартап-студии |
|-------------------------------------|---|--|------------------------------|--|---|--|
| Общий объём целевого рынка, TAM | \$114,5 млрд | \$26,6 млрд | \$6,15 млрд | \$89,2 млн | \$22,6 млн | \$97,7 млн |
| Доступный объём рынка, SAM | \$11,45 млрд | \$8,0 млрд | \$1,85 млрд | \$62,5 млн | \$6,38 млн | \$9,7 млн |
| Реально достижимый объём рынка, SOM | \$1,72 млрд | \$0,8 млрд | \$0,28 млрд | \$6,2 млн | \$0,6 млн | \$1,0 млн |

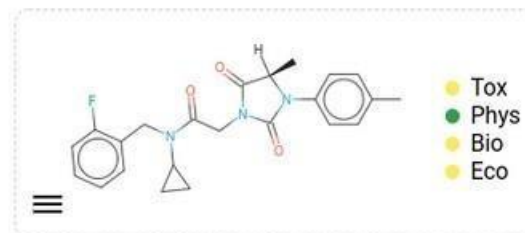
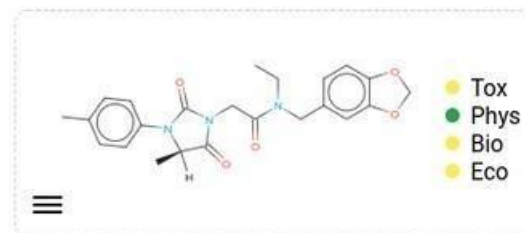
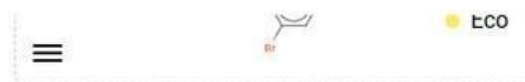
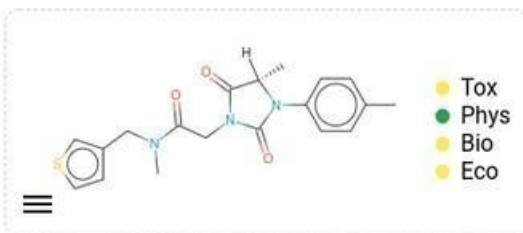
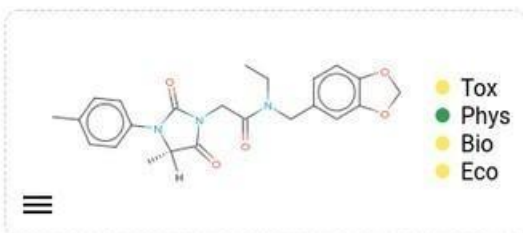
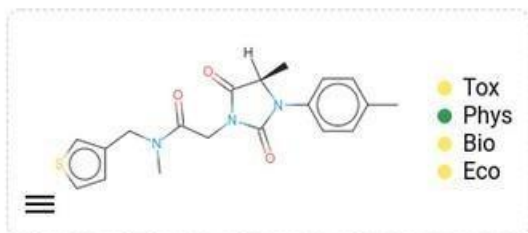
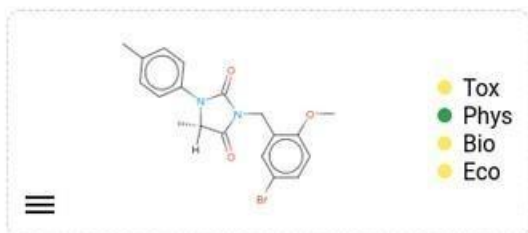
Пример №1:

Модуль базы данных.



SMILES, Brand Name, IUPAC Name, Vendor Code...

Search

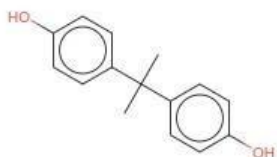
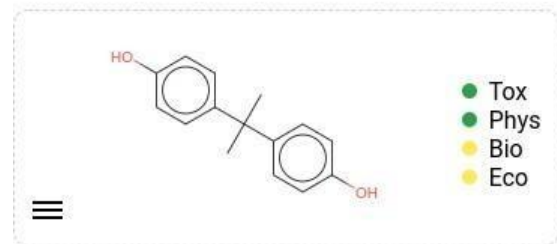


Пример №2:

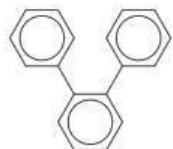
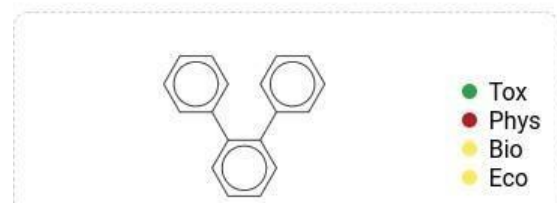
Модуль базы данных известных молекулярных структур с референсными ссылками на другие базы данных.



● Tox
● Phys
● Bio
● Eco



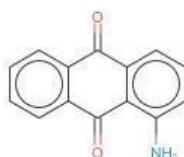
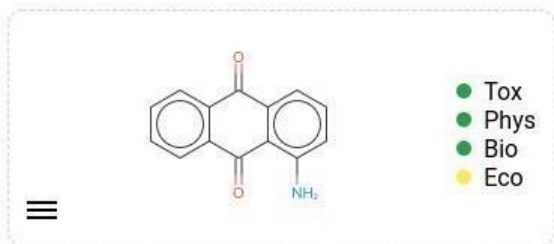
● Tox
● Phys
● Bio
● Eco



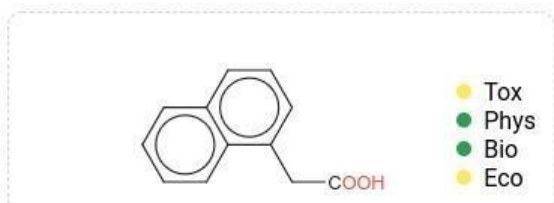
● Tox
● Phys
● Bio
● Eco



● Tox
● Phys
● Bio
● Eco



● Tox
● Phys
● Bio
● Eco



● Tox
● Phys
● Bio
● Eco

External DBs

| | |
|------------|---------------------------------|
| BindingDB | 50206424 ↗ |
| ChEMBL | CHEMBL220259 ↗ |
| SureChEMBL | SCHEMBL154124 ↗ |
| UNII | 9A1IU1C93L ↗ |
| ChEBI | CHEBI:87316 ↗ |
| DSSTOX | DTXSID7026520 ↗ |

Synonyms

87-02-5

j acid

7-amino-4-hydroxy-2-naphthalenesulfonic acid

7-amino-4-hydroxynaphthalene-2-sulfonic acid

isogamma acid

[OPEN TABLE](#)

Пример №3:

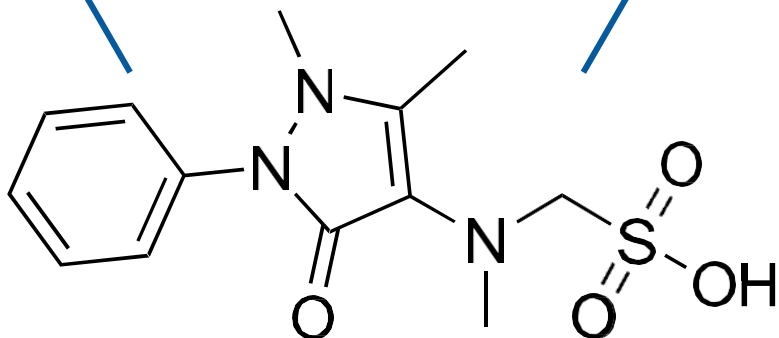
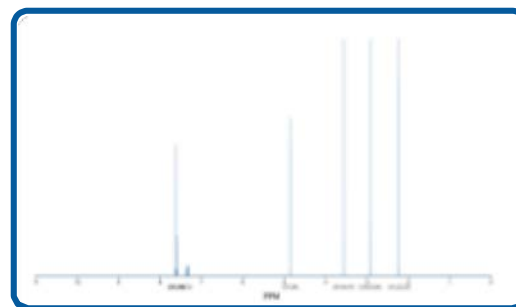
Модули расчета молекулярных свойств.

Drug-likeness

LogP = 0.2
QED = 0.84
TPSA = 92.92 Å²

Spectrum

¹H, ¹³C NMR



Physical

Melting point = 170 °C
Boiling point = 349 °C
Density = 1.34 g/cm³

Structure

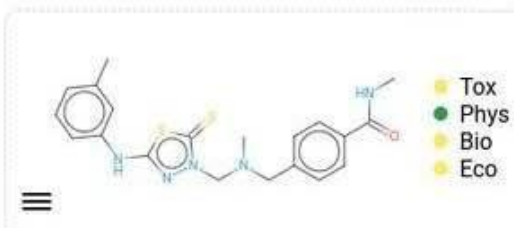
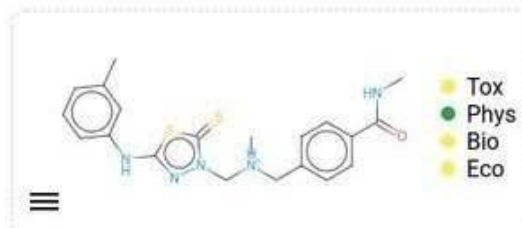
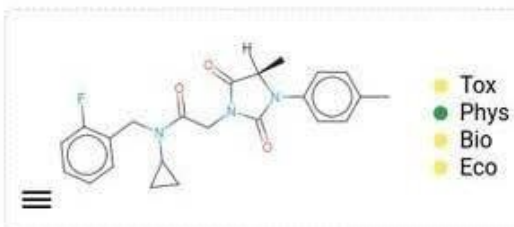
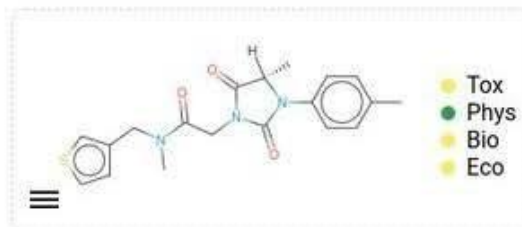
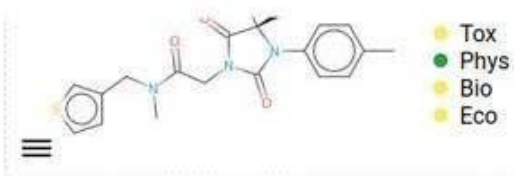
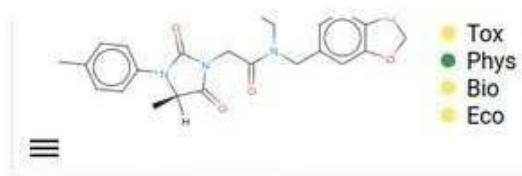
SMILES: Cc1c(N(C)CS(=O)(=O)O)c(=O)n(-c2ccccc2)n1C
IUPAC: [(1,5-dimethyl-3-oxo-2-phenylpyrazol-4-yl)-methylamino]methanesulfonic acid

- ✓ IUPASназвание
- ✓ ЯМР-спектр (H1)
- ✓ индексы подобия лекарству
- ✓ физико-химические свойства
- ✓ токсикологические свойства
- ✓ экологические свойства (фактор биоконцентрации)
- ✓ биологические свойства (ингибирование цитохромов)

Пример №4:

Модуль расчёта индивидуальных свойств соединений.

- IUPAC название
- ЯМР-спектр (H1)
- индексы подобия лекарству
- физико-химические свойства
- токсикологические свойства
- экологические свойства
- (фактор биоконцентрации)
- биологические свойства
- (ингибирование цитохромов)



methyl-3-(4-methylphenyl)-2,5-dioxoimidazolidin-1-yl]acetamide

OPEN TABLE

Bio

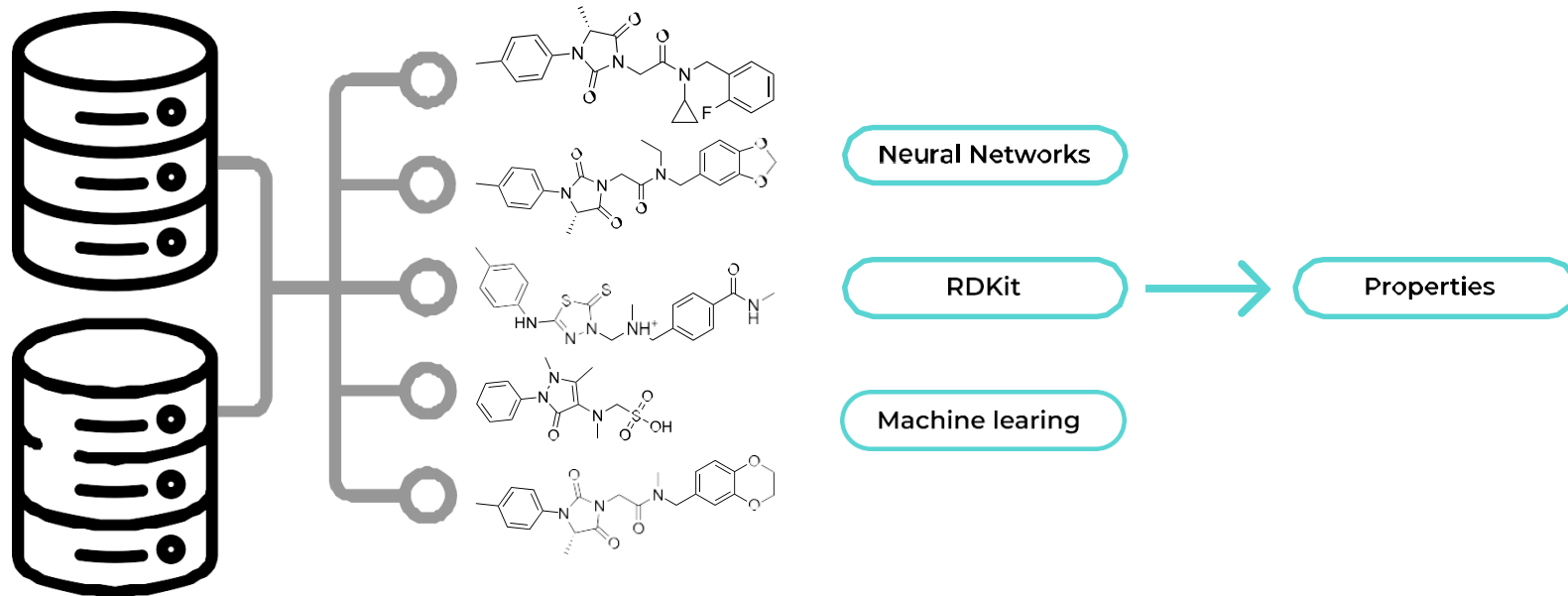
| | |
|---------|--------------|
| CYP1A2 | Noninhibitor |
| CYP2C19 | Inhibitor |
| CYP2C9 | Inhibitor |
| CYP2D6 | Noninhibitor |
| CYP3A4 | Inhibitor |

Physical

| | |
|-------------------------|---------------------------|
| Water Solubility | -4.4 LogS |
| Water Solubility (QSAR) | -4.653 ± 0.186 Log(mol/L) |
| Vapor Pressure | -6.4 log(mmHg) |

Пример №5:

Модули Искусственного Интеллекта для экспресс-оценки молекул-кандидатов.

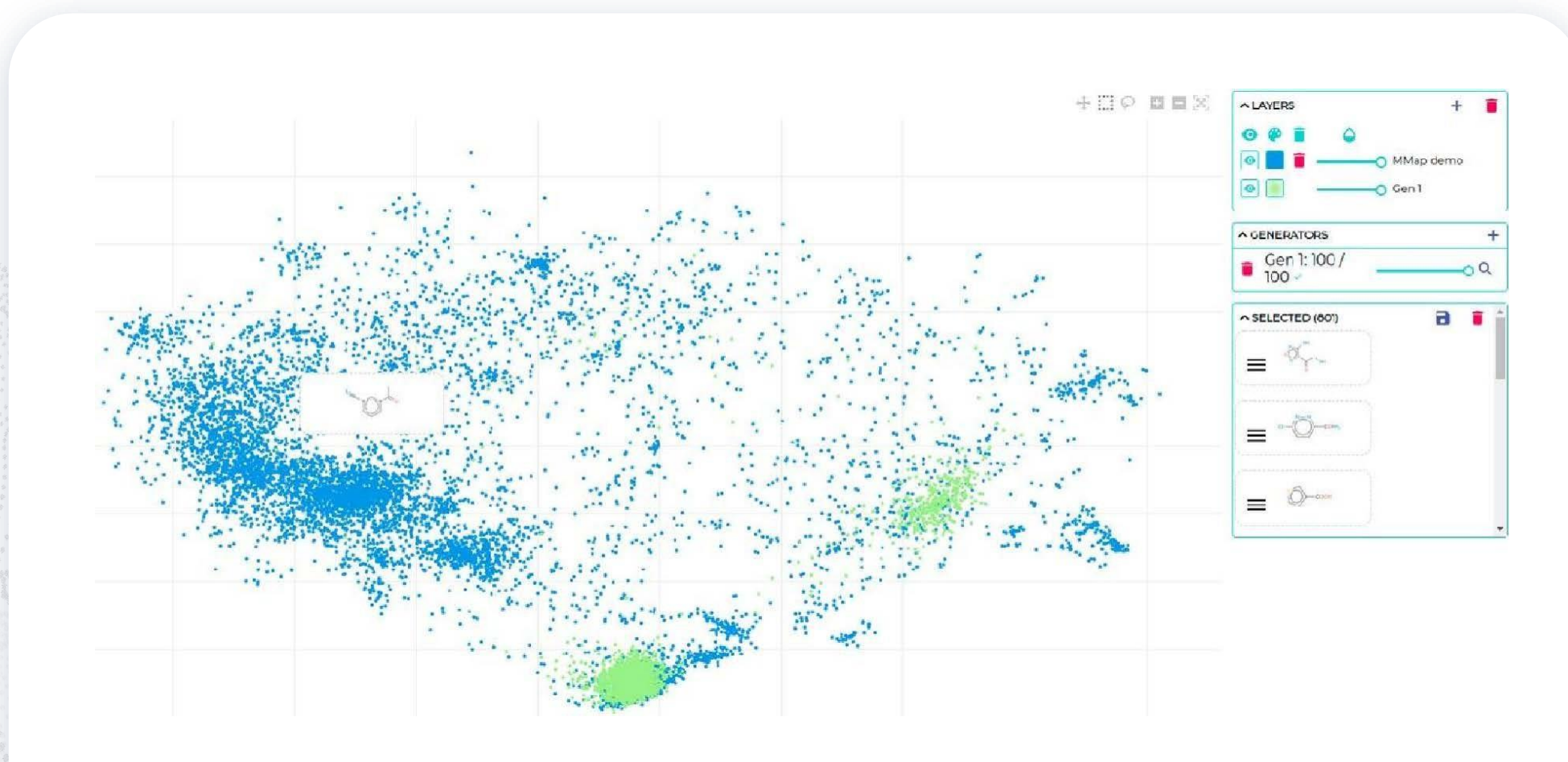


≥ 7 статей

опубликовано
в высокорейтинговых
журналах по
результатам НИКОР

ИНСТРУМЕНТЫ ДЛЯ ВИЗУАЛЬНОЙ НАВИГАЦИИ ПО ХИМИЧЕСКОМУ ПРОСТРАНСТВУ

- ✓ Анализ кластеров биоактивных соединений
- ✓ Анализ резких изменений активности (*activitycliffs*)



Извлечение химических **данных** из различных неструктурированных печатных источников (статьи, патенты, диссертации), их распознавание и анализирование.

- Public
- Molecules
- Reactions
- M-Map
- Individual
- SynWiki
- PDF2SMI
- RISM
- Admin

LOAD FILE

DOWNLOAD ALL

DOWNLOAD SELECTED

Scheme 1.

Table 1
Product ratios (IN:CB) of photolysis of 1-3 in various solvents

| Compound | Solvent | Product Ratios (IN:CB) |
|----------|--------------------|----------------------------|
| 1 | Benzene | 1:2.4 (26.1%:73.9%) |
| | MeOH | CB only |
| | CH ₃ CN | 1:1 (50%:50%) ^a |
| | Acetone | CB only (E:Z = 3.4:1) |
| | DMSO | CB only (E:Z = 2.2:1) |
| 2 | Benzene | CB only (E:Z = 3.7:1) |
| | MeOH | CB only (E:Z = 3.0:1) |
| | CH ₃ CN | 1:2.8 (26.4%:73.6%) |
| | Acetone | CB only |
| | DMSO | CB only |
| 3 | Benzene | CB only |
| | MeOH | CB only |
| | CH ₃ CN | CB only |
| | Acetone | CB only |
| | DMSO | CB only |

^a Normalized to 100%. Material balances of photoproducts including benzoic acid which is omitted in the table were in the range of 90-95% based on starting materials.

^b E-isomer only.

Source

Render

[R1]c1cc([R'])c2c(c1)C([R3])
([R2])CC2=O

18



*Используем глубокое
обучение*



*Предоставляем удобный
программный **интерфейс** для
создания плагинов сторонними
разработчиками*



*Полностью **облачная платформа** с
возможностью какразвертывания
на инфраструктуре заказчика,
так и на высокопроизводительных
кластерах длярешения масштабных
вычислительных задач*

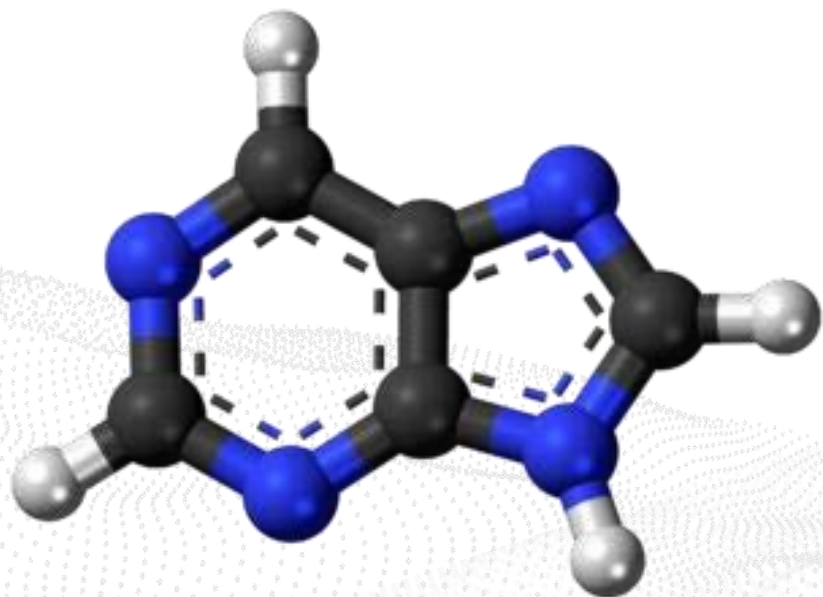
Мы планируем достичь:

Снижение времени на разработку
лекарств **на 3-4 года.**

Доли около **5-10% рынка** программного
обеспечения и услуг в области
вычислительной химии
и хемоинформатики.

Снижение затрат на разработку
лекарств **на 20%.**

Существует **встроенный экспериментальный модуль генерации** новых соединений с заданными свойствами из заранее заданных областей химического пространства.



Пример: Оптимизируем путем увеличения drug-likeness (QED), уменьшения токсичности и генерируем соединения преимущественно из области молекул схожих с производными пурина (северная область карты): оптимизируем LD50 Mouse + QED + точка на молекулярной карте.

ПРИМЕРЫ СОЗДАННЫХ СТРУКТУР

The screenshot displays the SynMap software interface, which is used for generating and analyzing chemical structures. The main area shows a grid of 12 generated structures, each with a legend indicating its properties: Tox (Toxicity), Phys (Physicochemical), Bio (Biological), and Eco (Ecological). The interface includes a sidebar with navigation options like Public, Molecules, Reactions, SynMap, Individual, SynWiki, PDF2SMILES, RISM, QSAR, and Admin. A top toolbar shows icons for zooming and a 'LAYERS' panel.

ГЕНЕРАТОР МОЖЕТ БЫТЬ ИСПОЛЬЗОВАН:

- ✓ Для оптимизации соединений-лидеров
- ✓ Для обхода патентных ограничений (!)
- ✓ Для поиска в заданной области химического пространства лучших кандидатов
- ✓ Для изучения структуры химического пространства

**БЛАГОДАРИМ
ЗА ВНИМАНИЕ!**



fcprii.rf