



РОССИЙСКО-БЕЛОРУССКИЕ ГЕНЕТИЧЕСКИЕ ТЕХНОЛОГИИ

Биомедицина и фармакология

Генетика животных

Сельскохозяйственная генетика
и биотехнология растений

Молекулярные биотехнологии

Биоинформатика

наука



И ТЕХНОЛОГИИ

Сибири

НАУКА И ТЕХНОЛОГИИ СИБИРИ

Выпуск 7 — Российско-белорусские генетические технологии
Ноябрь 2022 г.

Учредитель:

Федеральное государственное бюджетное учреждение
«Сибирское отделение Российской академии наук».
630090, Россия, Новосибирск,
проспект Академика Лаврентьева, дом 17.

Издатель:

Федеральное государственное бюджетное учреждение
«Сибирское отделение Российской академии наук».
630090, Россия, Новосибирск,
проспект Академика Лаврентьева, дом 17.

Главный редактор:

академик Валентин Николаевич Пармон.

Редакционный совет:

академики РАН Михаил Воевода, Николай Колчанов, Василий Фомин, Дмитрий Маркович, генеральный директор АО «Академпарк» Дмитрий Верховод, заместитель полномочного представителя Президента России в СФО Вадим Головкин, председатель Совета ректоров СФО профессор Николай Пустовой, заместитель председателя СО РАН д.ф.-м.н. Сергей Сверчков (ответственный за выпуск).

Редакционная группа:

Заместитель главного редактора Сергей Сверчков, Лариса Деева, Владимир Ларин, Андрей Соболевский, Татьяна Урбах, Любовь Батраева, Юлия Андреева.

Фото

авторов представленных материалов и из открытых источников.

Дизайн:

ФГБОУ ВО «Новосибирский государственный университет архитектуры, дизайна и искусств имени А.Д. Кряčkова» ректор Багрова Наталья, арт-директор Чешева Татьяна, дизайнеры: Теряева Анна, Перегудова Вероника, Юнг Виктория, Кирпичникова Снежана.

Свидетельство о регистрации СМИ ПИ № ФС 77–82311

от 03.12. 2021 г. выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

Адрес редакции: 630090, Россия, Новосибирск,
проспект Лаврентьева, 17, каб. № 224, тел.: 8 (383) 217-45-78,
e-mail: l.batraeva@sb-ras.ru

Отпечатано в ООО «Новосибирский издательский дом»
630048, г. Новосибирск, ул. Немировича-Данченко, 104
Тел.: (383) 299-29-80, e-mail: knjigosibirsk@yandex.ru
http://книгосибирск.рф/

Подписано в печать 14.11.2022

Бумага мелованная. Печать офсетная.
Тираж 800 экз. Свободная цена.

Перепечатка материалов только с письменного разрешения редакции.
Изданию присвоен номер ISSN: 2782-4969

© Сибирское отделение РАН, 2022





обращение главного редактора

4

экспертные статьи

6

биомедицина и фармакология

18

В номере

стр. 4 Обращение главного редактора академика В. Н. Пармона

стр. 6 Практическое использование генетических технологий в Беларуси. Опыт сотрудничества с Российской Федерацией

стр. 11 Научное братство

стр. 18 Метод выявления вероятности развития остеопороза с патологическими переломами

стр. 24 Фармакогенетический подход в оценке риска лекарственных осложнений при терапии шизофрении

стр. 28 Разработка способа количественной оценки генетической предрасположенности к развитию полигенных патологий

стр. 31 Использование полноэкзомного секвенирования для диагностики сложных случаев в педиатрии

стр. 35 Молекулярный тест для дооперационной диагностики узловых образований щитовидной железы

стр. 39 Углеродминеральные сорбенты на основе оксида алюминия СУМС-1

стр. 41 Прорывные разработки Института химической биологии и фундаментальной медицины СО РАН в области биофармацевтики и биомедицины

стр. 46 Способ выявления мутаций 2282del4, R501X, R2447X в гене филаггрина (FLG) при вульгарном ихтиозе и атопическом дерматите

стр. 48 Технология молекулярно-генетической диагностики семейной гиперхолестеринемии методом таргетного высокопроизводительного секвенирования

стр. 51 Способ выявления мутации p.L265P в гене MYD88

стр. 53 Технология диагностики и персонализированной терапии моногенных форм сахарного диабета

стр. 56 Технология лечения злокачественных новообразований «Каранакан»

стр. 64 Персонализированные технологии прогноза социально значимых заболеваний человека на основе молекулярно-генетических маркеров




генетика животных

68

стр. 68 Центр генетических ресурсов лабораторных животных — ключевая инфраструктура в импортозамещении модельных организмов для биоиспытаний

стр. 70 Использование генетических маркёров в животноводстве



сельскохозяйственная генетика и биотехнология растений

78

стр. 78 Молекулярно-цитогенетические методы, направленные на создание эффективной технологии получения гибридов тритикале и секалотритикум

стр. 83 Молекулярная генетика злаковых и плодовых культур для решения практических вопросов в сельском хозяйстве

стр. 87 Тест-система для экспресс-диагностики смешанных инфекций лесных древесных растений

стр. 96 Разработка и внедрение комплекса селекционно-генетических технологий и создание на их основе сортов зерновых культур, адаптированных к различным экологическим зонам

стр. 99 Применение биотехнологии соматического эмбриогенеза в качестве основы для программы MVF в России

стр. 103 Биотехнология получения посадочного материала перспективных для сибирского региона сортов голубики топяной и ее гибридов



молекулярные биотехнологии

110

стр. 110 Производство ферментных препаратов для нужд биотехнологии

стр. 114 Геномные биотехнологии для животноводства в Беларуси



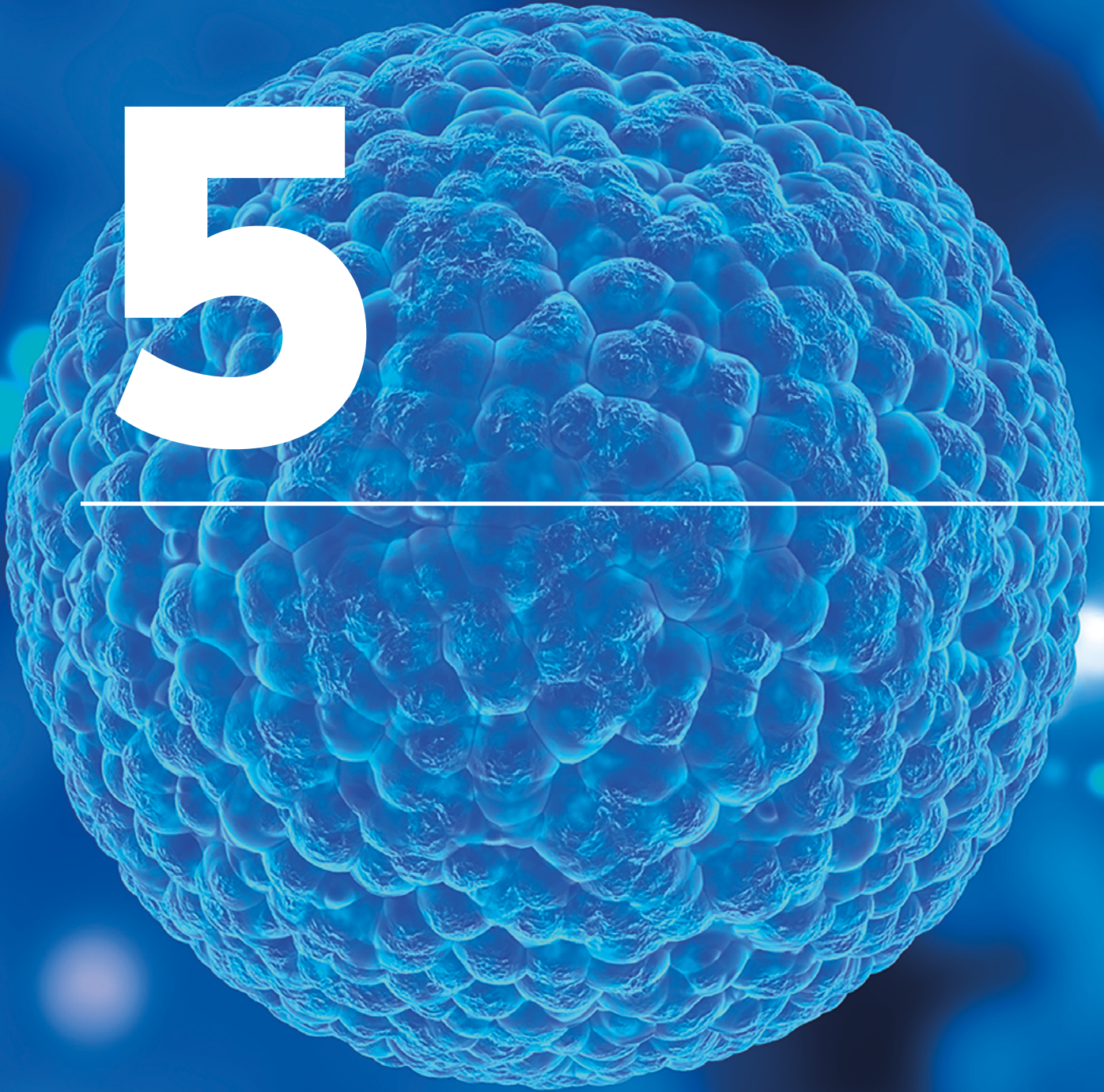
биоинформатика

122

стр. 122 ANDSsystem: когнитивная система для реконструкции и анализа графа знаний (генных сетей) на основе автоматического извлечения знаний из текстов научных публикаций, патентов и фактографических баз данных

стр. 126 ANDDigest — система семантического поиска информации в научных статьях и патентах на основе методов автоматического анализа текстов и машинного обучения

5



Биоинформатика

122

ANDSystem: когнитивная система для реконструкции и анализа графа знаний (генных сетей) на основе автоматического извлечения знаний из текстов научных публикаций, патентов и фактографических баз данных

126

ANDDigest — система семантического поиска информации в научных статьях и патентах на основе методов автоматического анализа текстов и машинного обучения

ANDSYSTEM: КОГНИТИВНАЯ СИСТЕМА ДЛЯ РЕКОНСТРУКЦИИ И АНАЛИЗА ГРАФА ЗНАНИЙ (ГЕННЫХ СЕТЕЙ) НА ОСНОВЕ АВТОМАТИЧЕСКОГО ИЗВЛЕЧЕНИЯ ЗНАНИЙ ИЗ ТЕКСТОВ НАУЧНЫХ ПУБЛИКАЦИЙ, ПАТЕНТОВ И ФАКТОГРАФИЧЕСКИХ БАЗ ДАННЫХ



Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», пр-т Академика Лаврентьева, 10, г. Новосибирск, 630090

Иванисенко Владимир Александрович

доцент, ведущий научный сотрудник, заведующий лабораторией, кандидат биологических наук
тел. (383) 363-49-80, +7-913-945-6649, salix@bionet.nsc.ru

Деменков Павел Сергеевич

доцент, научный сотрудник, кандидат технических наук
тел. (383) 363-49-80, +7-923-240-9542
demps@bionet.nsc.ru

Иванисенко Тимофей Владимирович

научный сотрудник
тел. (383) 363-4980, +7-913-923-0120, itv@bionet.nsc.ru

Колчанов Николай Александрович

академик РАН, профессор, научный руководитель, доктор биологических наук
тел. (383) 363-4980, +7-913-919-2364, kol@bionet.nsc.ru

Система с использованием методов искусственного интеллекта осуществляет автоматическую экстракцию знаний из мирового корпуса научных статей, патентов и баз данных и представляет их в удобном виде графа знаний для решения широкого круга задач.

Программно-информационная когнитивная система ANDSystem с использованием методов искусственного интеллекта автоматически извлекает знания из научных публикаций, международных патентов и фактографических баз данных, интегрирует их и удобно представляет в виде графа знаний (ассоциативных генных сетей). Вершинами графа ассоциативных генных сетей являются молекулярно-генетические объекты, лекарства, заболевания, фенотипические признаки и др. (всего 12 типов объектов). Ребра определяют взаимодействия между объектами, включая физические, химические, регуляторные и т. д. (всего 24 типа).

Реконструкция и анализ генных сетей представляют важнейшие этапы работ при решении как фундаментальных, так и прикладных задач. Известно, что формирование любых фенотипических признаков (молекулярно-генетических, биохимических, физиологических, морфологических и др.) обеспечивается с участием генных сетей — группами координированно функционирующих генов, взаимодействующих друг с другом через свои продукты (РНК, белки и метаболиты). ANDSystem кардинально ускоряет процесс решения широкого круга задач в различных областях биологических, биомедицинских и сельскохозяйственных наук и биотехнологий, обеспечивая быструю автоматизированную реконструкцию и глубокий анализ генных сетей на основе графа знаний, экстрагированных из мирового корпуса научных статей (проанализировано более 30 млн абстрактов из PubMed и более 7 млн полноразмерных статей из PMC), международных патентов и фактографических баз данных.

К настоящему времени число научных публикаций и патентов достигло таких объемов, которые невозможно обработать вручную. Указанные выше важные источники информации относятся к слабоструктурированным большим данным (big data). В связи с этим стали активно развиваться компьютерные методы автоматического извлечения знаний из текстов на естественном языке (text mining). Однако технологии text mining нашли практическое применение для реконструкции графа знаний только с появлением так называемых когнитивных систем, включающих полный цикл инженерии знаний — от извлечения фактов до их представления в виде графа знаний. Примерами когнитивных систем в области биологии и биомедицины служат Watson, Ingenuity Pathway Analysis (IPA), STRING, Pathway Studio. ANDSystem создана на основе оригинальной технологии text mining и не уступает существующим аналогам по полноте экстрагируемых знаний, но превосходит их по числу учитываемых типов взаимодействий и типов объектов, что позволяет более детально описывать изучаемые процессы. Основные модули ANDSystem показаны на рисунках 1, 2.

Свидетельства о государственной регистрации: свидетельство РФ о госрегистрации базы данных № 2008615929, свидетельство РФ о госрегистрации программы для ЭВМ № 2008615929.

Область применения

Молекулярная биология и генетика: интерпретация омиксных данных.

Биомедицина и фармакология: реконструкция молекулярно-генетических механизмов заболеваний и механизмов действия лекарств; идентификация новых фармакологических мишеней и биомаркеров; интерпретация генетических/эпигенетических вариаций в геноме человека, включая влияние мутаций на метаболизм лекарств, а также на целевые и нецелевые фармакологические эффекты лекарств, лекарство-лекарственные взаимодействия; приоритизация факторов риска заболеваний и диагностических маркеров для данного геномного статуса и др.

Сельскохозяйственная биология и растениеводство: изучение комплексных хозяйственно ценных признаков, таких как резистентность культур к заболеваниям и вредителям, устойчивость к воздействию абиотических факторов, урожайность и т. д.; выявление регуляторных и сигнальных путей ответа растения на конкретные условия роста, биотические и абиотические стрессы (высокая или низкая температура, засуха, засоленность, загрязнение почв металлами, реакция на удобрения, действие гормонов, элиситеров и т. д.); предсказание генов-кандидатов для генотипирования; поиск маркеров для маркер-ориентированной селекции; поиск кандидатов-мишеней для направленного воздействия веществами (в том числе внешними факторами) на растения для решения проблем раннего/равномерного всхода, лучшего вегетативного роста, эффективного поглощения питательных веществ и др.

Промышленная микробиология: идентификация биотехнологически значимых микроорганизмов и создание новых штаммов-продуцентов промышленных ферментов для пищевого и кормового производства, штаммов для биodeградации отходов, а также наработки аминокислот, биоспиртов и других ценных продуктов.

Уровень готовности технологии: TRL 7 — прототип системы продемонстрирован в реальных условиях эксплуатации.

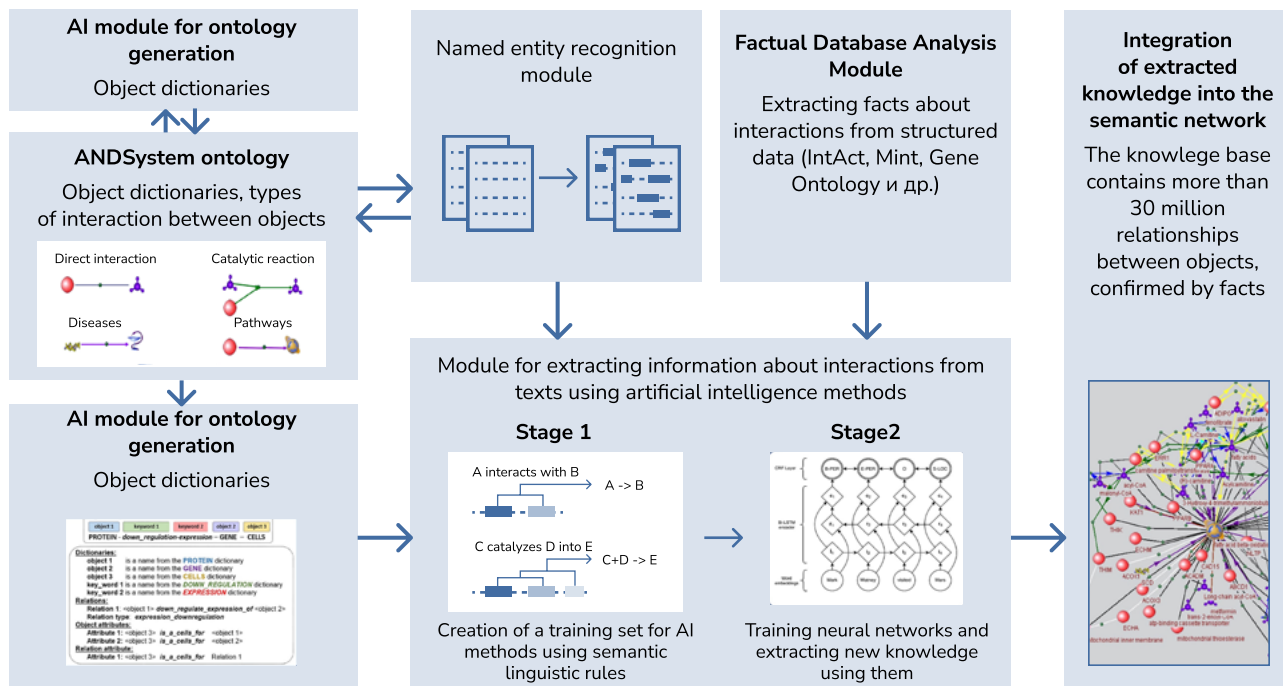
Предполагаемый интерес для внедрения: система может быть настроена на определенную предметную область и решение конкретных задач организации, а также установлена в ее

внутренней сети с разграничением прав доступа к информации в базе знаний и возможностью периодического обновления. В базу знаний ANDSystem могут быть экспортированы знания, полученные на основе собственных исследований, проводимых в организации, что

позволит их формализовать и интегрировать с мировыми литературными знаниями.

Запрос на индустриальное партнерство: внедрение технологии на предприятиях ИП, покупка лицензии ■

РИСУНОК 1. ОСНОВНЫЕ МОДУЛИ ANDSYSTEM



1) модуль онтологии ANDSystem, который включает подмодуль генерации онтологии предметной области с использованием методов искусственного интеллекта, подмодуль словарей объектов (сущностей) предметной области, а также базу знаний семантико-лингвистических правил (шаблонов) для извлечения знаний из текста о взаимодействии между объектами. Эти шаблоны созданы вручную экспертами и насчитывают более 10 000 штук;

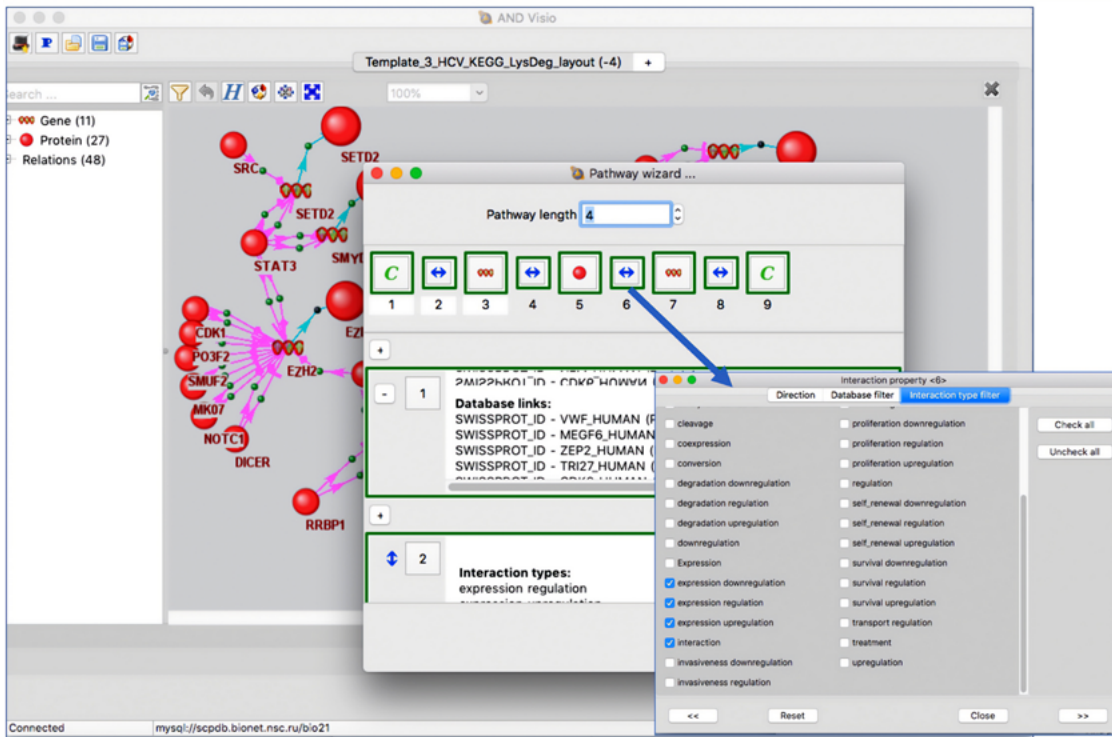
2) модуль предобработки текстов и распознавания имен объектов в текстах, в которых также используются методы искусственного интеллекта;

3) модуль извлечения данных из фактографических баз данных;

4) модуль извлечения знаний из текстов. Извлечение знаний из текстов осуществляется в два этапа. На первом этапе проводится извлечение знаний из размеченных текстов с применением семантико-лингвистических шаблонов. Второй этап включает обучение нейросетей и их применение для расширенного извлечения знаний. Для обучения используются знания, полученные с помощью семантико-лингвистических шаблонов;

5) информация, полученная с использованием модулей 3 и 4, интегрируется в базе знаний ANDCell.

РИСУНОК 2. ПРОГРАММА ANDVISIO



Программа предоставляет пользователю графический интерфейс для реконструкции генных сетей на основе обращения к базе знаний, их анализа и визуализации.



ANDDIGEST — СИСТЕМА СЕМАНТИЧЕСКОГО ПОИСКА ИНФОРМАЦИИ

В НАУЧНЫХ СТАТЬЯХ И ПАТЕНТАХ НА ОСНОВЕ МЕТОДОВ АВТОМАТИЧЕСКОГО АНАЛИЗА ТЕКСТОВ И МАШИННОГО ОБУЧЕНИЯ



Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», пр-т Академика Лаврентьева, 10, г. Новосибирск, 630090

Иванисенко Тимофей Владимирович

научный сотрудник
тел. (383) 363-49-80, +7-913-923-0120, itv@bionet.nsc.ru

Иванисенко Владимир Александрович

доцент, ведущий научный сотрудник, заведующий лабораторией, кандидат биологических наук
тел. (383) 363-49-80, +7-913-945-6649, salix@bionet.nsc.ru

Деменков Павел Сергеевич

доцент, научный сотрудник, кандидат технических наук
тел. (383) 363-49-80, +7-923-240-9542
demps@bionet.nsc.ru

Колчанов Николай Александрович

академик РАН, профессор, научный руководитель, доктор биологических наук
тел. (383) 363-49-80, +7-913-919-2364, kol@bionet.nsc.ru

Разработана система для поиска информации в мировых научных публикациях и патентах, которая с помощью методов искусственного интеллекта и фокусировки на конкретной предметной области значительно ускоряет работу с научно-технической информацией.

Программно-информационная система ANDDigest предназначена для интеллектуального поиска и анализа информации в мировых научных публикациях и международных патентах с использованием настраиваемой онтологии предметной области. В исходном варианте ANDDigest использует онтологию из когнитивной системы ANSystem, настроенную на области биомедицины и биотехнологии.

Поисковые запросы формируются с использованием тезаурусов онтологии, позволяющих осуществлять поиск на основе сущностей предметной области с учетом существующих синонимов, а также ключевых слов, вводимых пользователем вручную. Результаты поиска представляются размеченным набором текстов с объектами онтологии предметной области и в графическом виде — в виде графа семантических связей между объектами, представленными в найденных документах. Семантические связи между объектами, а также соответствие имени объекта в тексте его типу устанавливаются на основе методов искусственного интеллекта. Для каждого объекта генерируются тренды уровня интереса на основе динамики его упоминаний в публикациях и патентах. В системе реализована гибкая система фильтрации и сортировки найденных документов, включая фильтрацию по уровню статистиче-

ской значимости семантической связи между объектами, импакт-фактору научного журнала, дате публикаций и др. В России нет поисковых систем, обладающих аналогичными функциональными характеристиками. Применение в ANDDigest онтологии предметной области позволяет улучшить результаты поиска и существенно упростить формирование сложных запросов, а также обеспечивает возможность семантического анализа найденной информации, демонстрируя значительные преимущества по сравнению с существующими универсальными поисковыми системами (такими, как Google) при решении круга задач, связанных с анализом научных статей и патентов в рамках заданной тематики. Схема организации и веб-интерфейс ANDDigest приведены на рисунках 1–3. Прототип системы доступен по адресу <https://anddigest.sysbio.ru>.

Свидетельства о государственной регистрации: свидетельство о госрегистрации программы для ЭВМ № 2020660515; свидетельство о госрегистрации базы данных № 2020621827.

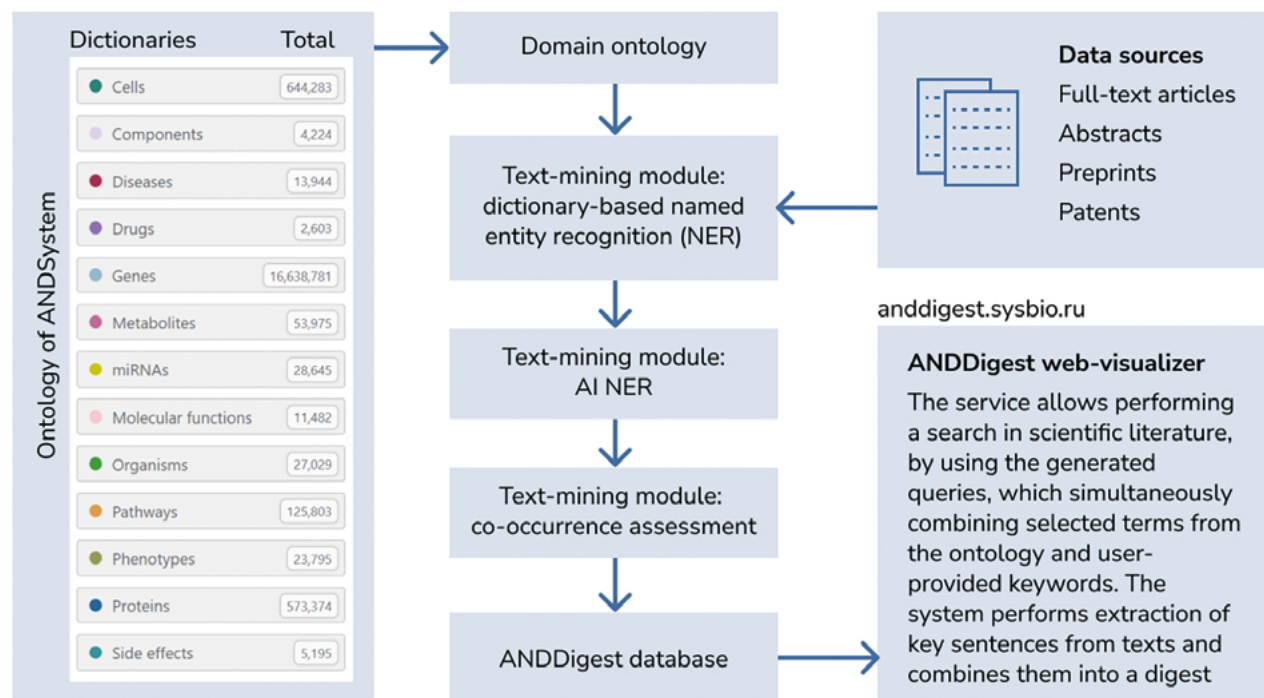
Области применения: базовая версия ANDDigest может применяться в области молекулярной и генетической биологии, биомедицине и фармакологии, а также в различных областях биотехнологий. Система является масштабируемой и может быть настроена на конкретную предметную область – например, промышленную микробиологию, сельскохозяйственную биологию и др.

Уровень готовности технологии: TRL 7 – прототип системы продемонстрирован в реальных условиях эксплуатации.

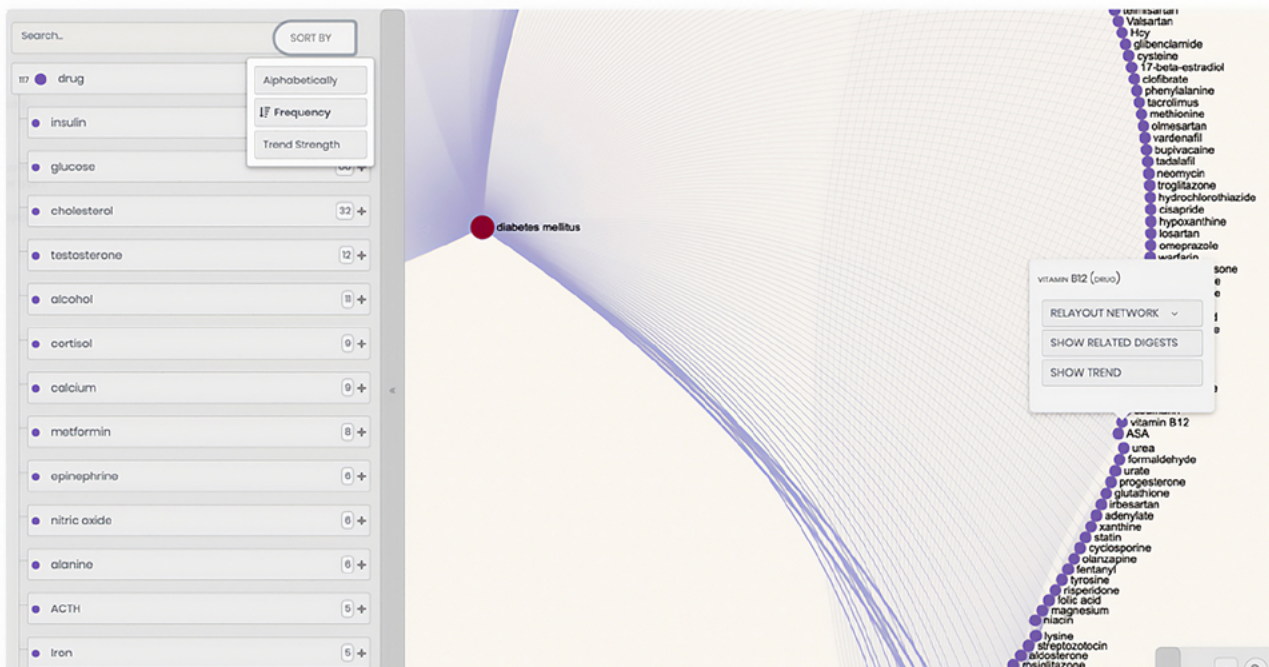
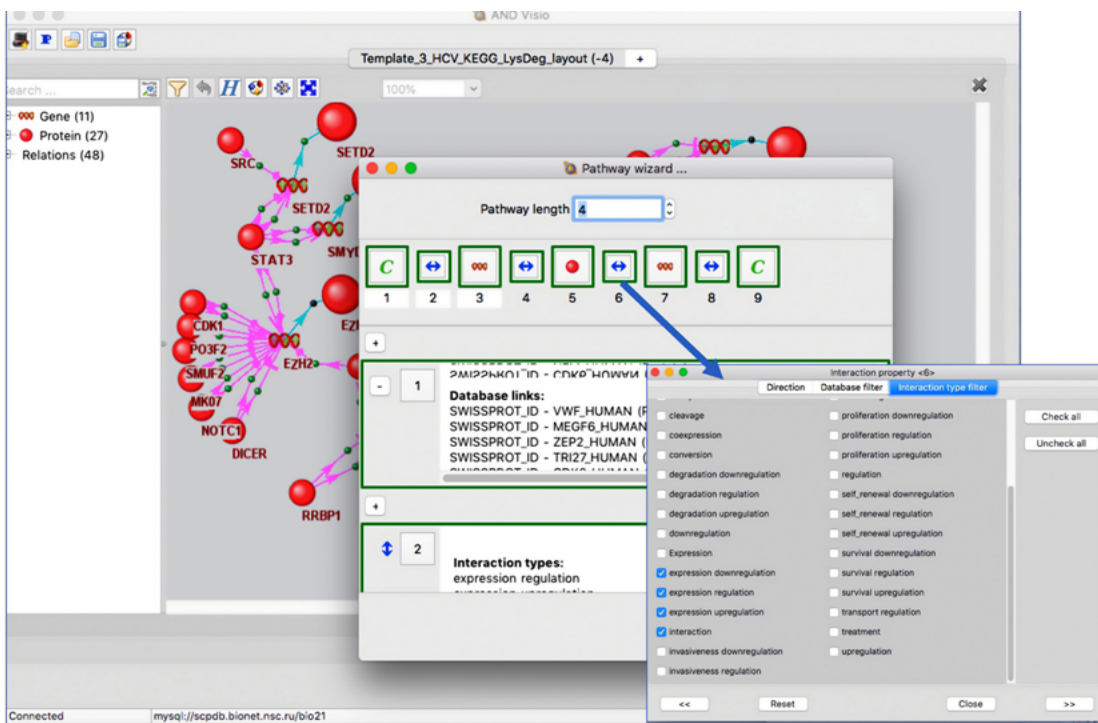
Предполагаемый интерес для внедрения: ANDDigest может быть применена в организациях, осуществляющих работу с научными публикациями и патентами. Использование ANDDigest позволит сократить время и улучшить качество поиска информации, а также автоматизировать мониторинг заданной предметной области.

Запрос на индустриальное партнерство: внедрение технологии на предприятиях ИП, покупка лицензии ■

СХЕМА ОРГАНИЗАЦИИ СИСТЕМЫ ANDDIGEST



ПОЛЬЗОВАТЕЛЬСКИЙ ВЕБ-ИНТЕРФЕЙС ANDDIGEST



ЦЕНТР УПРАВЛЕНИЯ ПРОЕКТАМИ СИБИРСКОГО ОТДЕЛЕНИЯ РАН: ЕДИНОЕ ОКНО В СФЕРУ ПОЛЕЗНОЙ НАУКИ



НАУЧНОЕ СОПРОВОЖДЕНИЕ КРУПНЫХ ИНДУСТРИАЛЬНЫХ ПРОЕКТОВ

Примеры:

- Междисциплинарная поддержка прокладки и эксплуатации газопровода «Сила Сибири - 2»
- Исследование биоразнообразия во всех дивизионах ПАО ГМК «Норникель»

ПРОЕКТЫ ПО СОЗДАНИЮ НОВЫХ ТЕХНОЛОГИЙ

Примеры:

- Утилизация промышленных, медицинских и коммунальных отходов
- Высокоэффективное сжигание топлив
- Перспективные виды транспорта



КОМПЛЕКСНАЯ ЭКСПЕРТИЗА ТЕХНОЛОГИЧЕСКИХ РЕШЕНИЙ ДО И ПОСЛЕ ВНЕДРЕНИЯ

Примеры:

- Состояние сооружений в Норильском промышленном районе
- Пригодность горнорудного оборудования
- Обоснованность штрафов за экологический ущерб

УЧАСТИЕ В ПРОГРАММАХ РАЗВИТИЯ ТЕРРИТОРИЙ

Примеры:

- Автоматизация управления городским транспортом
- Логистика медицинских учреждений и мест массового скопления людей
- Мини-ботанические сады для образовательных и детских учреждений



КОНТАКТЫ: начальник Деева Лариса Викторовна, pmc@sb-ras.ru, (383) 217-49-88

[ВЕРНУТЬСЯ К СОДЕРЖАНИЮ](#)

