

Научная статья

УДК 027.021:001.92:54(092):004.8

<https://doi.org/10.20913/2618-7515-2026-2-7>

Применение ИИ-технологий для популяризации научными библиотеками наследия выдающихся ученых: на примере трудов А. Е. Фаворского

Application of AI Technologies for Promoting the Heritage of Outstanding Scientists by Scientific Libraries: on the Example of A. E. Favorsky's Works

© **Кустова Елена Михайловна** ✉

старший библиограф, младший

научный сотрудник,

Центральная научная библиотека Федерального

исследовательского центра «Иркутский инсти-

тут химии им. А. Е. Фаворского СО РАН»,

ул. Фаворского, 1, Иркутск, 664033, Россия

ORCID: 0009-0009-0013-9604

e-mail: kustovahelen@mail.ru

© **Бараш Светлана Михайловна**

старший библиограф, младший

научный сотрудник,

Центральная научная библиотека Федераль-

ного исследовательского центра «Иркутский

институт химии им. А. Е. Фаворского СО РАН»,

ул. Фаворского, 1, Иркутск, 664033, Россия

ORCID: 0009-0004-7722-6574

e-mail: svm.barash@gmail.com

© **Трофимов Иван Леонидович**

зав. библиотекой, младший научный сотрудник,

Центральная научная библиотека Федераль-

ного исследовательского центра «Иркутский

институт химии им. А. Е. Фаворского СО РАН»,

ул. Фаворского, 1, Иркутск, 664033;

Государственная публичная научно-техническая

библиотека Сибирского отделения Российской

академии наук (ГПНТБ СО РАН),

ул. Восход, 15, Новосибирск, 630008, Россия

ORCID: 0000-0001-8654-6255

e-mail: t_john88@mail.ru

Kustova Elena Mikhailovna

Senior Bibliographer, Junior Researcher,

Central Scientific Library of the Federal Research

Center "Irkutsk Institute of Chemistry named after

A. E. Favorsky SB RAS",

1 Favorsky St., Irkutsk, 664033, Russia

ORCID: 0009-0009-0013-9604

e-mail: kustovahelen@mail.ru

Barash Svetlana Mikhailovna

Senior Bibliographer, Junior Researcher,

Central Scientific Library of the Federal Research

Center "Irkutsk Institute of Chemistry named after

A. E. Favorsky SB RAS"

1 Favorsky St., Irkutsk, 664033, Russia

ORCID: 0009-0004-7722-6574

e-mail: svm.barash@gmail.com

Trofimov Ivan Leonidovich

Head of the Library, Junior Researcher,

Central Scientific Library of the Federal Research

Center "Irkutsk Institute of Chemistry named

after A. E. Favorsky SB RAS",

1 Favorsky St., Irkutsk, 664033;

State Public Scientific Technological Library

of the Siberian Branch of the Russian Academy

of Sciences (SPSTL SB RAS),

15 Voskhod St. Novosibirsk, 630008, Russia

ORCID: 0000-0001-8654-6255

e-mail: t_john88@mail.ru

Аннотация. Целью статьи является представление новой формы сохранения и популяризации научного наследия ученых-химиков: создание интеллектуальной системы (ИС) на основе искусственного интеллекта (ИИ) для обработки химической литературы на примере публикаций академика А. Е. Фаворского.

Библиотеки применяют различные модели сохранения научного наследия и популяризации науки от лекций, выставок, мемориальных библиотек ученых, библиографических баз данных научных трудов до современных форматов, основанных на новейших генеративных моделях искусственного интеллекта. Авторы кратко описывают опыт Центральной научной библиотеки Федерального исследовательского центра «Иркутский институт химии им. А. Е. Фаворского СО РАН» по сохранению научного наследия ученых-химиков.

Развитие систем ИИ в области химии требует создания независимых массивов данных на основе проверенной информации, а также разработки интеллектуальных систем для машинного анализа текстов химической литературы. В статье обосновывается выбор трудов выдающегося химика-органика академика А. Е. Фаворского в качестве надежного первоисточника формирования эталонных обучающих выборок для оцифровки и дальнейшего распознавания химических графов в текстах работ ученого. Представлена характеристика создаваемых библиотекой ИИ-инструментов в целях использования их в работе по распознаванию химических текстов – применение технологий оптического распознавания химических структур (OCSR) для перевода молекулярных графов в машиночитаемый формат (SMILES).

Работа библиотеки по созданию современных интеллектуальных инструментов не только вводит в цифровой оборот исторические знания, но и пробуждает интерес молодых поколений к профессии ученого-химика.

Ключевые слова: научные библиотеки, оцифровка фонда, искусственный интеллект, большие языковые модели, оптическое распознавание химических структур (OCSR), формат SMILES, А. Е. Фаворский

Для цитирования: Кустова Е. М., Бараш С. М., Трофимов И. Л. Применение ИИ-технологий для популяризации научными библиотеками наследия выдающихся ученых: на примере трудов А. Е. Фаворского // Труды ГПНТБ СО РАН. 2026. № 2. С. 60–68. <https://doi.org/10.20913/2618-7515-2026-2-7>

Abstract. The purpose of this article is to present a new method for preserving and promoting the scientific heritage of chemists: creation of an intelligent system (IS), based on artificial intellect (AI), for processing chemical literature, on the example of publications by Academician A. E. Favorsky.

Libraries employ various models to preserve scientific heritage and popularize science – ranging from lectures, exhibitions, memorial libraries of scientists, and bibliographic databases of research publications to modern formats based on state-of-the-art generative AI patterns. The authors briefly describe the experience of the Central Scientific Library of the Federal Research Center Irkutsk Institute of Chemistry (Siberian Branch of the Russian Academy of Sciences) named after A. E. Favorsky in preserving the scientific heritage of chemists.

The development of AI systems in the field of chemistry requires the creation of independent datasets based on verified information, as well as the development of intelligent systems for the machine analysis of chemical literature. The article justifies the selection of the works of an eminent organic chemist – Academician A. E. Favorsky – as a reliable primary source for creating benchmark-training datasets intended for the digitization and subsequent recognition of chemical graphs within his texts. It presents AI tools, being developed by the library for chemical text recognition, specifically the application of Optical Chemical Structure Recognition (OCSR) technologies to convert molecular graphs into a machine-readable format (SMILES).

The library's work in creating modern intellectual tools not only brings historical knowledge into digital circulation but also sparks the interest of younger generations in the profession of a research chemist.

Keywords: scientific libraries, collection digitization, artificial intelligence, large language models, optical chemical structure recognition (OCSR), SMILES format, A. E. Favorsky

Citation: Kustova E. M., Barash S. M., Trofimov I. L. Application of AI Technologies for Promoting the Heritage of Outstanding Scientists by Scientific Libraries: on the Example of A. E. Favorsky' Works // Proceedings of SPSTL SB RAS. 2026. No. 2. P. 60–68. <https://doi.org/10.20913/2618-7515-2026-2-7>

Введение

Сохранение и продвижение научного наследия выдающихся ученых является важной составляющей процесса формирования общественного мировоззрения, обеспечения преемственности знаний, расширения образовательного кругозора молодежи, повышения научно-культурного уровня общества. В практическом решении этой задачи важное место занимают научные библиотеки Сибири. В статье О. Л. Лаврик описаны разные формы взаимодействия библиотек в сфере коммуникации науки и общества: создание и ведение различных баз данных трудов ученых (библиографических, полнотекстовых и т. п.), проведение массовых мероприятий (чтений, фестивалей, дней информации и др.), организация выставок литературы, ведение дискуссионных и иных площадок и др. [1]. Публикация М. А. Плешаковой и Д. В. Высоцкого посвящена обзору статей по организации и ведению мемориальных библиотек известных ученых, в том числе электронных (ЭМБ), выполняющих функции хранения культурной и научной памяти и продвижения научных знаний. Это, прежде всего, публикации методико-организационного плана по созданию ЭМБ сотрудников Государственной публичной научно-технической библиотеки Сибирского отделения Российской академии наук (ГПНТБ СО РАН), Библиотеки по естественным наукам РАН (БЕН РАН) и других учреждений (И. Г. Юдина, А. Ю. Бородихин, Т. Н. Ершова, М. Ч. Куулар, Л. П. Павлова, Е. Б. Виноградова, С. И. Зенкевич и др.) [2].

С развитием искусственного интеллекта (ИИ) появляются новые технологии и инструменты научной коммуникации: ИИ-ассистенты по поиску литературы, технологические программы по распознаванию научных текстов нейросетями, сгенерированные видеофильмы и художественные изображения (портреты ученых и др.), а также возникают новые формы распространения научных знаний, наряду с тематическими волонтерскими программами, научными телепередачами и т. п., – социальные медиа и онлайн-платформы, подкасты, научно-технические музеи и т. д. Специалисты БЕН РАН И. А. Митрошин и Н. А. Слащева называют научно-познавательный проект «Сайнс-арт» (Science Art, англ. «научное искусство»), где объединяется творчество, научные физико-биологические методы и ИТ-технологии – робототехника и др. Авторы считают, что посредством машинного обучения и современных GPT-продуктов можно создавать рекомендательные системы по поиску научной литературы и выявлять новые исследования [3]. Интересный формат сохранения историко-культурного наследия создан сотрудниками Национальной библиотеки Республики Саха (Якутия) – система «Истин» для обучения нейросетей якутскому языку и культуре, работающая на платформе VoiceNLRs (сбор голосовых данных для распознавания текстов и озвучивания речи

на якутском языке) [4]. В публикации В. К. Степанова представлены результаты научно-практической конференции «Искусственный интеллект в библиотечно-информационной деятельности», состоявшейся в Институте научной информации по общественным наукам РАН в мае 2025 г. В докладах российских и зарубежных коллег конференции (из Белоруссии, Казахстана, Китая и др.) приводятся примеры работы по внедрению ИИ-технологий в библиотечные процессы. Ряд докладов был посвящен описанию методик применения нейросетей для интеллектуальной обработки информации. Например, в докладе Зои Су (Tongfang Knowledge Network Technology) «Искусственный интеллект CNKI AI для максимального повышения эффективности работы исследователя» рассказывалось о нейроассистенте, встроенном в информационный массив одной из ведущих академических полнотекстовых платформ CNP [5].

Зарубежные библиотеки активно внедряют ИИ-технологии в библиотечное обслуживание пользователей. Например, о создании чат-ботов в помощь исследователям и студентам пишут Ю. Лаппалайнен и Н. Нараянан (библиотека Университета Зайеда, ОАЭ) [6], Ф. Боатен (США) [7] и др. Общие проблемы применения ИИ в библиотеках рассматриваются в статье L. S. Lo (Библиотека Университета Вирджинии, США) – конфиденциальность данных, вопросы равноправного доступа [8]. «Библиотеки уже не ассоциируются только с книгохранением, они успешно адаптируются к современной цифровой среде [9]. «Важным становится определение эффективных способов адаптации научных библиотек к сложноорганизованному современному социуму, изменению потоков научной информации и запросов самой науки» [10].

Актуальность задач сохранения и продвижения научного наследия известных ученых в работе отечественных научных библиотек поддерживается на государственном уровне. В «Стратегии научно-технологического развития Российской Федерации» (2024)¹ обозначена необходимость формирования эффективной системы коммуникации в области науки и технологий. Профессор Ю. Н. Столяров подчеркивает важную роль библиотек в претворении Стратегии в общественную и научную жизнь: он пишет о возможности создания Открытой библиотеки искусственного интеллекта и необходимости приобщения к ИИ-технологиям сотрудников библиотек, которые смогут создать «умные библиотеки» [11].

Центральная научная библиотека Федерального исследовательского центра «Иркутский институт химии им. А. Е. Фаворского СО РАН» (ЦНБ ФИЦ ИРИХ СО РАН) на протяжении нескольких лет

¹ Стратегия научно-технологического развития Российской Федерации : утв. Указом Президента Рос. Федерации от 28 февр. 2024 г. № 145 // Гарант.ру : информ.-правовой портал. URL: <https://www.garant.ru/products/ipo/prime/doc/408518353/> (дата обращения: 13.04.2026).

реализует задачи сохранения исторического научного наследия региона, продвижения научных знаний посредством создания уникальных информационно-образовательных ресурсов: научных и научно-популярных статей об ученых; электронных библиографических и полнотекстовых баз данных (ЭБ трудов геологов С. И. Шермана, И. А. Кобеляцкого и др.); ЭМБ известных сибирских ученых (например, байкаловеда профессора М. М. Кожова). Информационные ресурсы, созданные библиотекой, представлены на сайте ЦНБ².

С 2025 г. в ЦНБ реализуется научно-исследовательская работа (НИР) по разработке локальных нейросетевых архитектур для мультимодального извлечения знаний (OCSR) и семантического анализа химической литературы. На совещании в Кремле по вопросам развития ИИ в апреле 2026 г. Президент РФ В. В. Путин отметил: «Искусственный интеллект наряду с цифровыми платформами и автономными системами формирует принципиально иной облик <...> социальной сферы, образования [и научной сферы – прим. авт.] <...>, и только при наличии собственных моделей мы сможем уверенно двигаться вперед, гарантировать безопасность и обороноспособность, что особенно важно, оставаться на переднем крае научной и инженерной мысли»³. В связи с этим ЦНБ выступает не просто хранилищем физических носителей, но и провайдером верифицированных данных для обучения интеллектуальных систем (ИС).

Тематика НИР библиотеки предполагает распознавание оцифрованных текстов публикаций выдающихся ученых-химиков и извлечение сложных химических формул с помощью алгоритмов ИИ для создания научных тезаурусов. Сотрудники ЦНБ рассматривают роль библиотеки как «опоры проверенной достоверной информации» [12]. Система ИИ, которая может распознавать химические формулы в научных статьях и переводить их в понятный компьютеру язык, поможет изучению архивов старых научных журналов. Труды ученых по химии представляют особую сложность из-за отсутствия общепринятых стандартов, связанных с оформлением химических формул и структур⁴. В статье описан начальный этап создания в ЦНБ ИС, реализующей двухпоточную архитектуру поиска и распознавания химических текстов на основе ИИ-технологий. Выбор трудов А. Е. Фаворского как информационно-эксперимен-

тальной базы ИС обосновывается значительным вкладом ученого в становление и развитие иркутской химической науки, внушительным образовательным тезаурусом знаний, основанным на его научных публикациях, практическом вкладе в развитие народного хозяйства страны. Работа носит комплексный научно-исследовательский характер и включает биобиблиографическое исследование фонда библиотеки для выявления трудов ученого и подготовки публикаций к оцифровке, а также – практико-ориентированное информационно-технологическое направление по созданию ИС для поиска и распознавания химической литературы на основе ИИ-технологий.

Цель исследования – сохранение и продвижение научного наследия известных ученых-химиков, чья деятельность связана со становлением и развитием академической химической науки Восточной Сибири, через создание ИС на основе ИИ для обработки химической литературы (на примере публикаций академика А. Е. Фаворского). ИС будет способствовать доступности химических научных текстов для исследований, обучения специалистов-химиков и студентов, повышения знаний старшеклассников, интересующихся химией. Выявленная информация о деятельности и научных трудах академика А. Е. Фаворского будет размещена на мемориальной странице ученого на сайте ЦНБ ФИЦ ИРИХ СО РАН.

А. Е. Фаворский: биография. Научная школа Фаворского в Иркутске

Алексей Евграфович Фаворский (1860–1945) – выдающийся химик-органик, академик АН СССР (1929), специалист в области явлений изомеризации и полимеризации ацетиленовых углеводородов. Лауреат Сталинской премии первой степени. Наследие ученого представляет собой огромную историческую ценность и формирует корпус химических текстов, подходящий для обучения алгоритмов интеллектуального поиска.

Окончив Петербургский университет, Алексей Евграфович работал в нем на протяжении почти 60 лет, заведовал кафедрой органической химии (рис. 1). Он преподавал химию и в других учебных заведениях, в том числе на Высших женских (Бестужевских) курсах (1900–1918 гг.), где организовал первую научно-исследовательскую лабораторию в женском учебном заведении [14] (рис. 2).

А. Е. Фаворский был одним из создателей Государственного института прикладной химии (ГИПХ) и Института органической химии (ИОХ) АН СССР в Москве (в 1934–1938 гг. – директор института)⁵.

² Центральная научная библиотека ИРИХ СО РАН: сайт. URL: <https://csl.isc.irk.ru/> (дата обращения: 13.04.2026).

³ Совещание по вопросам развития технологий искусственного интеллекта. Президент провел в Кремле совещание по вопросам развития технологий искусственного интеллекта // Президент России : сайт. URL: <http://kremlin.ru/events/president/news/79525> (дата обращения: 13.04.2026). Дата публикации: 10.04.2026.

⁴ Исследователи научили нейронную сеть распознавать химические формулы из научных статей // Научная Россия : сайт. URL: <https://scientificrussia.ru/articles/issledovateli-naucili-nejronnuu-set-raspoznavat-himicheskie-formuly-iz-naucnyh-statej> (дата обращения: 13.03.2026). Дата публикации: 14.02.2022.

⁵ Левченков С. И. Фаворский Алексей Евграфович (3 марта 1860 г. – 8 августа 1945 г.) // Краткий очерк истории химии : учеб. пособие. URL: <https://www.physchem.chimfak.sfedu.ru/Source/History/Persones/Favorsky.html> (дата обращения: 16.03.2026).



Рис. 1. А. Е. Фаворский в год защиты докторской диссертации, 1895 г. Фото Т. Фаворской
Fig. 1. A. E. Favorsky in the year of defending his doctoral dissertation, 1895. Photo by T. Favorskaya

В период Великой Отечественной войны (1941–1945 гг.) А. Е. Фаворский продолжал вести научную работу в эвакуации (Казахстан, курорт «Боровое»), вместе с М. Ф. Шостаковским исследовал синтез виниловых эфиров. Полученный полимер – знаменитый «бальзам Фаворского – Шостаковского» («Винилин») был очень востребован в годы войны для лечения ожогов и ран. В своей научной деятельности А. Е. Фаворский всегда стремился внедрить научные достижения в промышленности, народном хозяйстве, фармакологии: производство органического стекла и синтетических каучуков, синтез полимеров – пластмассы, смолы, из которых вырабатываются лаки, пленки, клей, искусственный шелк и т. д. [14; 15].

Последователями школы А. Е. Фаворского были директор Иркутского института химии СО РАН: член-корреспондент АН СССР М. Ф. Шостаковский (1957–1970), академики М. Г. Воронков (1970–1994) и Б. А. Трофимов (1994–2015). Научная школа Б. А. Трофимова (химия ацетилена) развивается и в настоящее время⁶. В 2000 г. ИРИХ СО РАН было присвоено имя А. Е. Фаворского, в фойе учреждения установлен его бюст. Улица, на которой расположен институт, также названа в честь ученого.

Создание корпуса химических данных на основе оцифрованных трудов А. Е. Фаворского

В ходе исследования фондов ЦНБ были обнаружены сборники избранных трудов академика А. Е. Фаворского (1934, 1940, 1961 годов изданий),

⁶ Научная школа академика Б. А. Трофимова // Федеральный исследовательский центр «Иркутский институт химии им. А. Е. Фаворского Сибирского отделения Российской академии наук»: сайт. URL: <https://www.irkinstchem.ru/vedushchie-nauchnye-shkoly/shkola-akademika-b-a-trofimova> (дата обращения: 28.05.2026).

а также уникальный массив отечественной и иностранной периодики, в которой публиковался ученый:

«Доклады АН СССР» (с 1922 г. по настоящее время); «Журнал Русского физико-химического общества» (1873–1930 гг.) и «Журнал общей химии» (1931–2010 гг.); Journal fur Praktische Chemie (экземпляры 1958–1968 гг.)⁷; Justus Liebig's Annalen der Chemie (1832–1997 гг.); Comptes Rendus Hebdomadaires des Seances de L' Academie des Sciences (Париж, 1912–1965 гг.); Chemische Berichte (Berichte der Deutschen Chemischen Gesellschaft). Зарегистрированы экземпляры с 1868 по 1997 г.

Помимо карточного и электронного каталогов ЦНБ (АБИС ИРБИС64), были обследованы электронные каталоги крупных библиотек России для выявления трудов ученого и уточнения их библиографического описания. Найдены некоторые авторские свидетельства А. Е. Фаворского и их оцифрованные тексты – в БД eLIBRARY, в Российском государственном архиве г. Самары (РГА г. Самары). В результате исследования был выявлен ряд трудностей, препятствующих прямой машинной обработке изданий:

низкое качество физических носителей: пожелтевшая бумага, пятна, рукописные пометки, заломы (выпуски отдельных томов журналов были туго переплетены) и сложность верстки; отсутствие единого стандарта представления химических формул; отсутствие современного планетарного сканера; нехватка вычислительных мощностей.

Чтобы превратить бумажные фонды в активный цифровой центр знаний, необходимо не просто «отсканировать», а переосмыслить информацию. Для этих целей специалистами и библиографами ЦНБ было начато создание ИС на основе ИИ-технологий.

Технологии распознавания химической литературы

Выбор трудов академика А. Е. Фаворского для реализации первого этапа НИР обоснован не только их историко-культурной значимостью, но и сложностью представленных в них химических структур, что позволяет отработать методы распознавания химических данных. Интересными и важными являются формулы и химические реакции, выработанные А. Е. Фаворским и получившие названия: «перегруппировка Фаворского», «правила Фаворского», «реакция Фаворского».

⁷ Немецкие журналы по химии поступили в фонд ЦНБ в 1960-х гг. – были перевезены из Германии после Победы СССР в Великой Отечественной войне (1941–1945) в порядке репарации. На экземплярах сохранилась печать Берлинского университета: «Technisch-Chemisches Institut der Technischen Hochschule Berlin».



Рис. 2. Лаборатория органической химии А. Е. Фаворского и К. И. Дебу на Высших женских курсах в Санкт-Петербурге (1878–1918 гг.). Фото из открытых источников

Fig. 2. The organic chemistry laboratory of A. E. Favorsky and K. I. Debu at the Higher Courses for Women in St. Petersburg (1878–1918). Photo from open sources

В 1887 г. ученый открыл изомеризацию ацетиленовых углеводородов под влиянием спиртового раствора щелочи (ацетилен-алленовая перегруппировка), которая стала общим методом синтеза ацетиленовых и диеновых углеводородов («перегруппировка Фаворского»), а затем сформулировал закономерности протекания этих реакций («правила Фаворского»). В 1905 г. – реакцию получения третичных ацетиленовых спиртов конденсацией ацетиленовых углеводородов с карбонильными соединениями в присутствии безводного порошкообразного гидроксида калия («реакция Фаворского») [16; 17].

С точки зрения информатики и технологий компьютерного зрения (ComputerVision), химические схемы и формулы представляют собой нелинейные двухмерные графы. Традиционные системы оптического распознавания символов (OCR, такие как Tesseract), алгоритмы которых настроены на линейный текст, непригодны для анализа химической литературы. Базовый OCR-модуль воспринимает бензольное кольцо как хаотичный набор букв «О» или геометрических примитивов, двойные связи интерпретируются как знаки равенства, а стереохимические клинья (указывающие пространственное расположение атомов) считаются как типографский шум.

Для преодоления этой проблемы в рамках НИР внедряются технологии оптического распознавания химических структур – OCSR (Optical Chemical Structure Recognition). Архитектура разрабатываемой системы использует много-ступенчатый подход. Первоначально растровое изображение страницы старинного журнала анализируется моделью детекции объектов

(например, YOLO или специализированными Decimer и rxnScribe). Эти модели классифицируют области страницы, выделяя зоны, где находятся формулы. Вырезанные фрагменты с формулами передаются в мультимодальные модели (Vision-LanguageModels), адаптированные специально для химии (наподобие архитектуры MolScribe, Qwen-VL или DeepSeek-OCR).

Задача этого ИИ-конвейера – не просто «сфотографировать» молекулу, а понять ее топологию, валентность связей и осуществить семантический перевод растрового молекулярного графа в одномерную строковую номенклатуру SMILES (Simplified Molecular Input Line Entry System) (рис. 3).

Рассмотрим скан страницы документа с молекулярным графом процесса образования α -хлоризопрена (при t кип. 100°C) при помощи реакции с малеиновым ангидридом (Фаворский А. Е., Фаворская Т. А. Ацетилен-аллен-диеновая ступенчатая перегруппировка галоидгидринов (1961)) [18].

Рабочий процесс распознавания молекулярного графа (OCSR):

1. Сегментация формулы-реакции с использованием rxnScribe.
2. Разбивка сложной реакции на отдельные формулы для преобразования.
3. Результат перевода формулы в машиночитаемый формат SMILES с использованием MolScribe: O=C1C=CC(=O)O1.
4. Проверка на валидность с использованием Rdkit: Формула – C4H2O3.
5. Затем полученная запись в формате SMILES сохраняется в базе данных (СУБД PostgreSQL) для дальнейшей работы.

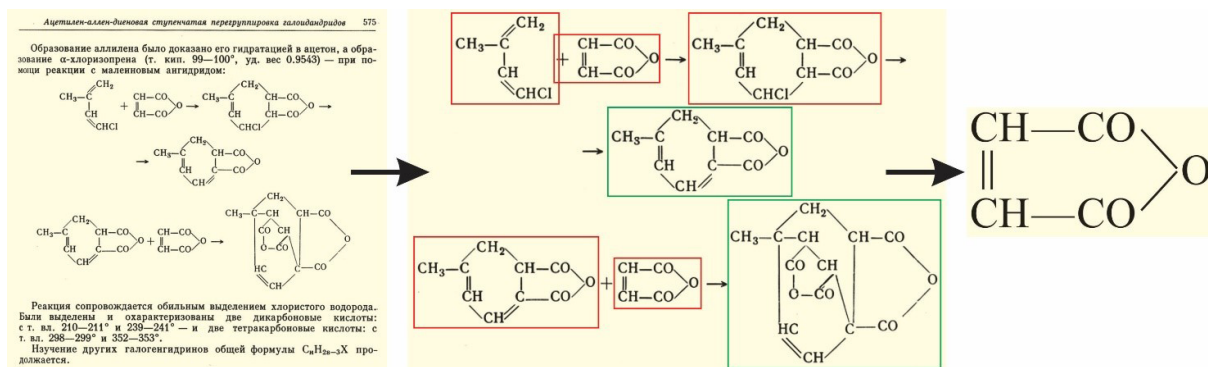


Рис. 3. Рабочий процесс распознавания молекулярного графа (OCSR)
Fig. 3. Optical Chemical Structure Recognition Workflow of the molecular graph

Таким образом разрабатываемый в рамках НИР ИИ-ассистент сможет не только находить публикации ученых, но и осуществлять поиск по приведенным в них химическим реакциям. Управление столь разнородным историческим фондом в цифровую эпоху потребовало от специалистов ЦНБ разработки концепции гибридного поиска – Dual Stream Retrieval (двухпоточный поиск, интеграция текста и OCSR).

Программный комплекс ИС, применяемый в работе, принципиально отличается от стандартных библиотечных информационно-поисковых систем, функционирующих на алгоритмах поиска по метаданным (автор, название, ключевые слова). Главной задачей ИС является не просто нахождение документа в фонде, а извлечение растрового изображения формулы, анализ ее топологии, валентности связей и последующий перевод (трансляция) в машиночитаемый строковый формат химической номенклатуры SMILES. ИИ здесь не выступает как внешняя надстройка для поиска, а является технологическим ядром системы. Подробное техническое описание архитектуры ИС, качества распознавания молекулярных графов и программной реализации выходит за рамки нашей публикации и будет представлено в отдельной работе.

На текущем этапе разрабатываемая ИС представляет собой набор программных алгоритмов, интегрированных с локальным сервером библиотеки, на котором запускаются и тестируются ИИ-модели.

Заключение

Работа по выявлению трудов академика А. Е. Фаворского в фонде ЦНБ ФИЦ ИРИХ СО РАН позволила не только исследовать ценный массив исторической периодики, но и заложить научно-методический фундамент для цифровой трансформации библиотеки. Установлены сборники избранных трудов ученого разных лет, в РГА г. Самары обнаружены отдельные тематические

издания и источники⁸, освещающие результаты научно-исследовательской деятельности Алексея Евграфовича. В редком фонде библиотеки выявлены ценные периодические издания по химии – отечественные и иностранные XIX–XX вв., где опубликованы труды ученого, начата работа по оцифровке отдельных изданий. Проводится разработка и апробация ИС на основе ИИ-технологий по распознаванию оцифрованных текстов химической литературы.

Впервые в работе академической библиотеки научно-исследовательская деятельность ученого-химика рассматривается как тестовый корпус для обучения мультимодальных нейросетей, распознавания химических молекулярных графов (OCSR) и трансляции их в форматы SMILES. Созданная ИС может быть использована как основа для формирования мемориального ресурса и для обучения студентов-химиков.

Статья выполнена в рамках НИОКТР (НИР) «Разработка локальных нейросетевых архитектур для мультимодального извлечения знаний (OCSR) и семантического анализа химической литературы с использованием больших языковых моделей», № 125122515025-5

Авторы прочитали и одобрили окончательный вариант рукописи.

Вклад авторов

Е. М. Кустова: историко-научный анализ деятельности академика А. Е. Фаворского, аналитический обзор литературы, выборка тестовых графов; написание текста статьи, кроме раздела «Технологии распознавания химической литературы», обсуждение и научное редактирование рукописи.

⁸ Марченко О. И., Покровская Л. Ю. Фаворский Алексей Евграфович (1860–1945) // Российский государственный архив в г. Самаре : сайт. URL: http://rga-samara.ru/activity/publications/smi/articles/345/?sphrase_id=36488 (дата обращения: 11.03.2026). [Представлены архивные источники о деятельности А. Е. Фаворского].

С. М. Бараш: биобиблиографический поиск, историко-научный анализ деятельности академика А. Е. Фаворского, выборка тестовых графов; участие в обсуждении и редактировании текста статьи.

И. Л. Трофимов: разработка и описание интеллектуальной системы, написание текста раздела «Технологии распознавания химической литературы», обсуждение и научное редактирование рукописи.

Конфликт интересов. Авторы заявляют об отсутствии конфликтов интересов, требующих раскрытия в этой статье.

Список литературы

1. Лаврик О. Л. Библиотека как участник процесса использования результатов научных исследований // Библиотековедение. 2024. Т. 73, № 4. С. 295–308. DOI: <https://doi.org/10.25281/0869-608X-2024-73-4-295-308>
2. Плешакова М. А., Высоцкий Д. В. Библиотеки как хранилища и популяризаторы культурной памяти и наследия: обзор практик создания и использования мемориальных библиотек ученых // Книга. Чтение. Медиа среда. 2025. Т. 3, № 4. С. 415–423. DOI: <https://doi.org/10.20913/BRM-3-4-9>
3. Митрошин И. А., Слащева Н. А. Новые горизонты популяризации науки в библиотеках // Научные и технические библиотеки. 2024. № 10. С. 141–156. DOI: <https://doi.org/10.33186/1027-3689-2024-10-141-156>
4. Климов С. С., Бурнашева А. Г., Федоров И. К., Баторов А. Р., Борисов В. Б., Максимова С. В. Искусственный интеллект в деятельности Национальной библиотеки Республики Саха (Якутия) // Библиотековедение. 2025. Т. 74, № 4. С. 343–360. DOI: <https://doi.org/10.25281/0869-608X-2025-74-4-343-360>
5. Степанов В. К. Искусственный интеллект в библиотечно-информационной деятельности: год спустя // Научные и технические библиотеки. 2025. № 12. С. 127–143. DOI: <https://doi.org/10.33186/1027-3689-2025-12-127-143>
6. Lappalainen Y., Narayanan N. Aisha: a custom AI library chatbot using the ChatGPT API // Journal of Web Librarianship. 2023. Vol. 17, no. 3. P. 37–58. DOI: <https://doi.org/10.1080/19322909.2023.2221477>
7. Boateng F. The transformative potential of generative AI in academic library access services: opportunities and challenges // Information Services and Use. 2025. Vol. 45, no. 1/2. P. 140–147. DOI: <https://doi.org/10.1177/18758789251332800>
8. Lo L. S. AI policies across the globe: implications and recommendations for libraries // IFLA Journal. 2023. Vol. 49, no. 4. P. 645–649. DOI: <https://doi.org/10.1177/03400352231196172>
9. Лизунова И. В., Пшеничная Е. В. Внедрение искусственного интеллекта и развитие научных коммуникаций (из опыта международной научно-практической конференции «LIBWAY-2024: Наука, технологии и информация в библиотеках») // Библиография и книговедение. 2024. № 3. С. 127–139. DOI: <https://doi.org/10.25281/2411-2305-2024-3-127-139>
10. Трескова П. П., Артемьева Е. Б. Научные библиотеки как ресурс научно-технологического развития России //

Труды ГПНТБ СО РАН. 2023. № 2. С. 25–28. DOI: <https://doi.org/10.20913/2618-7515-25-28>

11. Столяров Ю. Н. Искусственный интеллект и книжная библиотечная отрасль: направления разработки проблемы // Научные и технические библиотеки. 2022. № 1. С. 17–34. DOI: <https://doi.org/10.33186/1027-3689-2022-1-17-34>
12. Трофимов И. Л., Кустова Е. М., Бараш С. М., Буряк А. В., Филатова В. Н. Роль библиотек в эпоху искусственного интеллекта: аналитический взгляд на новые ИИ-технологии и их применение // Современные технологии. Системный анализ. Моделирование. 2025. № 3. С. 94–104.
13. Фаворская Т. А. Алексей Евграфович Фаворский, 1860–1945. Ленинград : Наука, Ленингр. отд-ние, 1980. 253 с.
14. Фаворская Т. А. Фаворские. Жизнь семьи университетского профессора. 1890–1953 : воспоминания. Санкт-Петербург : Изд-во С.-Петерб. ун-та, 2019. 781 с.
15. Шостаковский М. Ф. Алексей Евграфович Фаворский. Москва ; Ленинград : Изд-во АН СССР, 1948. 99 с.
16. Волков В. А., Вонский Е. В., Кузнецова Г. И. Фаворский Алексей Евграфович (3.03.1860 – 8.08.1945) // Химики : биограф. справ. Киев, 1984. С. 509–510.
17. Данилкина Н. А., Васильева А. А., Балова И. А. Научное наследие А. Е. Фаворского в современной органической химии: прототропная ацетилен-алленовая изомеризация и реакция «ацетиленовой молнии» // Успехи химии. 2020. Т. 89, № 1. С. 125–171. DOI: <https://doi.org/10.1070/RCR4902>
18. Фаворский А. Е. Избранные труды. Москва ; Ленинград : Изд-во Акад. наук СССР, 1961. 790 с.

References

1. Lavrik OL (2024) Library as a participant in the utilization of research results. *Bibliotekovedenie* 73 (4): 295–308. (In Russ.). DOI: <https://doi.org/10.25281/0869-608X-2024-73-4-295-308>
2. Pleshakova MA and Vysotskii DV (2025) Libraries as repositories and popularizers of cultural memory and heritage: practices of creating and using memorial libraries of scientists. *Kniga. Chtenie. Mediasreda* 3 (4): 415–423. (In Russ.). DOI: <https://doi.org/10.20913/BRM-3-4-9>
3. Mitroshin IA and Slashcheva NA (2024) New horizons of science popularization in libraries. *Nauchnye i tekhnicheskie biblioteki* 10: 141–156. (In Russ.). DOI: <https://doi.org/10.33186/1027-3689-2024-10-141-156>
4. Klimov SS, Burnasheva AG, Fedorov IK, Batorov AR, Borisov VB and Maksimova SV (2025) Artificial intelligence in the activities of the National Library of the Republic of Sakha (Yakutia). *Bibliotekovedenie* 74 (4): 343–360. (In Russ.). DOI: <https://doi.org/10.25281/0869-608X-2025-74-4-343-360>
5. Stepanov VK (2025) Artificial intelligence in library and information activities: one year later. *Nauchnye i tekhnicheskie biblioteki* 12: 127–143. (In Russ.). DOI: <https://doi.org/10.33186/1027-3689-2025-12-127-143>
6. Lappalainen Y and Narayanan N (2023) Aisha: a custom AI library chatbot using the ChatGPT API. *Journal of Web Librarianship* 17 (3): 37–58. DOI: <https://doi.org/10.1080/19322909.2023.2221477>
7. Boateng F (2025) The transformative potential of generative AI in academic library access services: opportunities and challenges. *Information Services and Use* 45 (1/2): 140–147. DOI: <https://doi.org/10.1177/18758789251332800>

8. Lo LS (2023) AI policies across the globe: implications and recommendations for libraries. *IFLA Journal* 49 (4): 645–649. DOI: <https://doi.org/10.1177/03400352231196172>
9. Lizunova IV and Pshenichnaya EV (2024) AI introduction and the scientific communication development (from the experience of the international scientific and practical conference "LIBWAY-2024: Science, technologies and information in libraries"). *Bibliografiya i knigovedenie* 3: 127–139. (In Russ.). DOI: <https://doi.org/10.25281/2411-2305-2024-3-127-139>
10. Treskova PP and Artemyeva EB (2023) Scientific libraries as a resource for the scientific and technological development of Russia. *Trudy GPNTB SO RAN* 2: 25–28. (In Russ.). DOI: <https://doi.org/10.20913/2618-7515-25-28>
11. Stolyarov YuN (2022) Artificial intelligence and the book library industry: vectors for problem development. *Nauchnye i tekhnicheskie biblioteki* 1: 17–34. (In Russ.). DOI: <https://doi.org/10.33186/1027-3689-2022-1-17-34>
12. Trofimov IL, Kustova EM, Barash SM, Buryak AV and Filatova VN (2025) The role of libraries in the era of artificial intelligence: an analytical look at new AI technologies and their applications. *Sovremennye tekhnologii. Sistemnyi analiz. Modelirovanie* 3: 94–104. (In Russ.).
13. Favorskaya TA (1980) Aleksey Evgrafovich Favorsky, 1860–1945. Leningrad: Nauka, Leningr. otd-nie. (In Russ.).
14. Favorskaya TA (2019) The Favorskys. The life of a university professor's family. 1890–1953: memoirs. Saint Petersburg: Izd-vo S.-Peterb. un-ta. (In Russ.).
15. Shostakovskii MF (1948) Aleksey Evgrafovich Favorsky. Moscow, Leningrad: Izd-vo AN SSSR. (In Russ.).
16. Volkov VA, Vonskii EV and Kuznetsova GI (1984) Aleksey Evgrafovich Favorsky (3.03.1860 – 8.08.1945). *Khimiki: biogr. sprav.* Kiev, pp. 509–510. (In Russ.).
17. Danilkina NA, Vasil'eva AA and Balova IA (2020) A. E. Favorskii's scientific legacy in modern organic chemistry: prototropic acetylene-allene isomerization and the acetylene zipper reaction. *Uspekhi khimii* 89 (1): 125–171. (In Russ.). DOI: <https://doi.org/10.1070/RCR4902>
18. Favorsky AE (1961). Selected works. Moscow, Leningrad: Izd-vo Akad. nauk SSSR. (In Russ.).

Статья поступила в редакцию / Received 27.04.2026

Получена после доработки / Revised 09.06.2026

Принята для публикации / Accepted 01.08.2026