

# Методы оценки применимости цифровых информационных ресурсов в науковедческих исследованиях (на примере научных идей Г.Ф. Гаузе)

Д.Е. Прокудин, Г.С. Левит

**Аннотация**—Развитие информационного общества сопровождается постепенным переводом в цифровую форму научной информации, представленной традиционно в печатном виде. Для её хранения создаются различные научные цифровые ресурсы с доступом из глобальной информационной сети. Это позволяет эффективно пользоваться научной информацией в целях оперативного проведения научных исследований, в том числе и науковедческих. Настоящее исследование направлено на изучение возможности использования цифровых информационных ресурсов для применения в науковедческих исследованиях. Предлагается метод отбора ресурсов на примере исследования влияния научных идей Г.Ф. Гаузе на развитие науки. На основе выбранных поисковых критериев выявляются особенности применения поисковых инструментов информационных ресурсов. Делаются выводы о дальнейшем развитии методов сбора и анализа получаемых из информационных ресурсов данных. При этом рассматриваются как мировые, так и российские информационные ресурсы. Статья отражает дальнейшее развитие тематического исследования, первые результаты которого обнародованы в материалах международной научной конференции *Internet and Modern Society (IMS-2017)* и XIX Всероссийской научной конференции «Научный сервис в сети Интернет».

**Ключевые слова**—научная информация, цифровые информационные ресурсы, методы отбора, критерии отбора, цифровизация научного наследия, Г.Ф. Гаузе.

## I. ВВЕДЕНИЕ

Неуклонный рост информации, связанный с развитием человеческого общества и объективным накоплением знаний, влияет и на увеличение объёмов научной информации. Основным механизмом увеличения объёмов информации является переход к её цифровому представлению, что влечёт ускорение процессов генерации информации, увеличение объёмов хранимой в электронной форме информации с использованием информационно-коммуникационных технологий. Развитие в последние 15-20 лет коммуникационных технологий позволило обеспечить глобальный и

широкий доступ к накопленным массивам цифровой информации. На сегодня наиболее активно развиваются полностью цифровые формы представления научной информации [7, 8, 10, 20, 21, 25].

Цифровые научные ресурсы, как правило, содержат информацию о каждой отдельной научной статье (или другой форме публикации) и описание статьи содержит все необходимые метаданные. К тому же, такие ресурсы обладают мощным поисковым механизмом, что позволяет быстро и эффективно осуществлять поиск необходимой информации по заданным критериям. Большинство цифровых научных ресурсов имеют доступ из сети Интернет, что решает как проблему доступности, так и проблему оперативного распространения результатов научных исследований [14, 15, 22]. Помимо этого, аккумуляция научных текстов в цифровой форме позволяет успешно решать задачу долговременного их хранения [5, 6, 9, 12, 16, 17]. Тем самым, решается проблема сохранения мирового культурного наследия, частью которого является и научное наследие. Это относится как к раритетным и уникальным изданиям, представляющим историческую ценность, так и к современным изданиям, которые потенциально являются объектами научного и культурного наследия для будущих поколений.

Современное развитие науки в условиях информатизации характеризуется интенсификацией как исследовательских процессов, так и процессов распространения результатов научных исследований. Поэтому для современных учёных и исследователей важнейшей задачей является оперативный и широкий по охвату источников доступ к постоянно прирастающему объёму научного знания. В этих условиях необходимо оценить возможности использования для полноценных исследований различных по типу и доступности цифровых научных ресурсов, а также выработать эффективные методы их использования.

## II. ЦИФРОВЫЕ НАУЧНЫЕ РЕСУРСЫ

В настоящее время наиболее развитыми являются технологии, направленные на аккумуляцию научного знания, его оперативного представления широкой общественности в цифровой форме с предоставлением удобных механизмов поиска. Как правило, они представляют собой базы данных метаданных публикаций с файловым хранилищем полных текстов

Статья получена 20 октября 2017.

Д.Е. Прокудин, докт. филос. наук, доцент, СПбГУ, аналитик, Университет ИТМО, Россия; (e-mail: hogben.young@gmail.com).

Г.С. Левит, докт. биол. наук, доцент, Университет ИТМО, Россия (e-mail: gslevit@corp.ifmo.ru).

соответствующих публикаций с доступом к поисковому механизму в этих базах данных из сети Интернет. Среди множества подобного рода ресурсов можно провести градацию по различным параметрам, что позволит определить возможности их использования в науковедческих исследованиях.

Можно выделить следующие типы цифровых научных информационных ресурсов:

1. Электронные каталоги библиотек.

2. Цифровые репозитории научной информации, которые получили широкое распространение с развитием парадигмы «Открытой науки» через инициативы Open Archives и Open Access. Их можно разделить на следующие основные категории [23]:

- институциональные, т.е. созданные и администрируемые различными организациями, в которых проводится научная деятельность (университеты, научные центры, научные институты и т.д.);

- тематические, предназначенные для аккумуляции результатов исследований по определённым научным направлениям;

- репозитории научных сообществ, которые содержат публикации, осуществлённые в рамках деятельности соответствующих сообществ.

3. Каталоги или регистры открытых репозиторий. На сегодняшний день широко известны два таких каталога – OpenDOAR (<http://www.opendoar.org>) и Registry of Open Access Repositories – ROAR (<http://roar.eprints.org>), которые содержат наиболее полную информацию об открытых репозиториях. На сайтах этих каталогов есть возможность поиска по контенту зарегистрированных в них репозиторий.

4. Электронные научные журналы (или электронные версии печатных журналов) и тематические коллекции научных сообществ, включающие в себя, в том числе, сборники научных статей и труды научных конференций. Как правило, для публикуемых статей на сайтах электронных журналов размещаются все необходимые метаданные, которые могут быть проиндексированы как информационно-поисковыми системами (в том числе и Google Scholar), так и быть переданы во внешние информационные системы. Для этого в основном используются либо импорт метаданных в определённом формате, либо автоматизированная передача метаданных по протоколу OAI-PMH. В этом случае данные представляются в формате Dublin Core, который является одним из общепринятых стандартов [11, 24].

5. Каталоги электронных журналов. К ним, в основном, относятся каталоги журналов открытого доступа (Open Access), например:

Directory of Open Access Journals (DOAJ, <http://doaj.org>) – каталог научных журналов открытого доступа. Создан в 2002 году совместно и при поддержке Лундского университета, Шведской национальной библиотеки и Королевской библиотеки (Стокгольм). В каталог включены рецензируемые электронные научные журналы свободного доступа. В настоящее время DOAJ предоставляет поиск на уровне статей из более чем 9 тыс. журналов (около 2,4 млн статей);

Directory of Open Access scholarly Resources (ROAD, <http://road.issn.org>) – сервис, созданный Международным центром ISSN при поддержке Сектора коммуникации и информации ЮНЕСКО в 2013 году. В этом каталоге содержатся сведения более чем о 18 тыс. журналах открытого доступа и других свободно распространяемых серийных изданиях.

6. Агрегаторы представляют собой информационные системы, аккумулирующие метаданные статей и других научных текстов, размещённых на соответствующих ресурсах, зарегистрированных в агрегаторах. Основной механизм пополнения баз данных агрегаторов состоит в автоматизации обмена метаданными по соответствующим протоколам, основным из которых является OAI-PMH. Описание любого объекта, размещённого в агрегаторе, помимо метаданных содержит и ссылку на источник предоставляемых данных. При этом в случае ресурсов открытого доступа по ссылке будет доступен и полный текст. Как пример можно привести один из самых мощных мировых агрегаторов The OAIster® database (<http://oaiSTER.worldcat.org>), который содержит более 30 млн. записей мета данных публикаций более чем 1,5 тыс. организаций-участников, предоставляющих свои публикации по принципу «открытого доступа» [18, 19].

7. Определённый интерес представляют реферативные базы данных, которые размещают метаданные в основном из периодических изданий (в основном это авторитетные научные журналы). Общепризнанными лидерами среди них являются Web of Science (сервис компании Thomson Reuters) и Scopus (сервис компании Elsevier). Однако, специфика этих баз данных такова, что они содержат публикации из определённых изданий, которые должны удовлетворять достаточно жёстким критериям, оставляя за бортом подавляющее большинство мировой научной периодики.

8. Научные социальные сети (например, Academia.edu и ResearchGate), которые весьма активно используются в исследовательской среде для свободного распространения научной информации, а также в целях научной коммуникации. Их наполнение происходит посредством самостоятельной ручной загрузки текстов и их метаданных через личные учётные записи пользователей. Поэтому информация, аккумулируемая в этих ресурсах, является разрозненной и фрагментарной.

Также все цифровые научные ресурсы можно разделить по доступности в них размещаемых текстов на обще доступные или открытые (как ресурсы, поддерживающие инициативу «Открытые архивы» и «Открытый доступ»), коммерческие (в основном платформы коммерческих издательств, а также реферативные базы WoS и Scopus) и смешанные, в которых есть как документы с открытым доступом, так и предоставляемые на платной основе (например, НЭБ).

Немаловажным аспектом для использования цифровых ресурсов в научных целях является степень детализации предоставляемой информации – в ресурсах могут содержаться либо только библиографические записи, либо библиографические записи могут сопровождаться метаданными, либо к этому ещё есть

наличие полных текстов публикаций.

Систематизировать цифровые электронные ресурсы можно также и по типу размещаемых в них документов: монографии, периодические и сериальные издания, диссертации, сборники статей, материалы конференций, патенты, отчёты, депонированные рукописи, статьи и т.д.

Важной характеристикой для исследователей является степень охвата научных направлений, представленных в цифровых ресурсах: узкоспециализированные), ориентированные на несколько научных направлений) или политематические, которые не ограничены по тематике научных направлений.

Также большой интерес в плане поиска научной информации представляют как информационно-поисковые системы общего назначения (например, Google, Yandex), так и специализированные ИПС (к которым на сегодня можно отнести Академию Google).

### III. ЗАДАЧА ОТБОРА НАУЧНЫХ ЦИФРОВЫХ РЕСУРСОВ ДЛЯ ИССЛЕДОВАНИЯ СТЕПЕНИ ВЛИЯНИЯ ИДЕЙ Г. Ф. ГАУЗЕ НА РАЗВИТИЕ МИРОВОЙ НАУКИ

Целью настоящего исследования является оценка возможности использования различных научных цифровых ресурсов для выявления влияния Гаузе на развитие различных областей науки и их отбор для дальнейшего качественного исследования.

Георгий Францевич Гаузе (1910-1986) был одним из самых влиятельных советских экологов и эволюционных биологов [1, 2]. Он считается классиком экспериментальной экологии как в России, так и на Западе. Известен он, прежде всего, как автор закона (принципа) конкурентного исключения, указавшего на невозможность устойчивого сосуществования двух (или более) видов в одной экологической нише при наличии факторов, лимитирующих рост их численности [13]. Самая известная книга Гаузе «The Struggle for Existence» (Борьба за Существование) была опубликована в США, при активной поддержке американского эколога, теоретического биолога и генетика Раймонда Перля (1879-1940) на английском языке и признана «одним из самых влиятельных произведений в теоретической и экспериментальной экологии» [4]. Теоретические математические модели динамики популяций, конкурирующих за определенный пищевой ресурс, были в 1920-х предложены итальянцем, почетным членом АН СССР Вито Вольтерра (1860-1940) и американцем австрийского происхождения Альфредом Лоткой (1880-1949). Теоретическая близость исследований Гаузе, Лотки, Вольтерра и Перля, интенсивные контакты между Гаузе и его западными партнерами, делают их вполне сравнимыми фигурами и позволяют, в перспективе, оценить сравнительную динамику их упоминаний в научной литературе.

Менее известен на Западе, вклад Гаузе в становление медицины в СССР и открытие им, совместно с супругой Марией Георгиевной Бражниковой, первого в СССР лечебного антибиотика Грамицидина С в 1942 г. [3]. Применение этого антибиотика в военно-полевых условиях способствовало заживлению гнойных ран

красноармейцев. За эту работу Гаузе, вместе с соавторами, был награжден Сталинской премией в 1946 г.

Еще менее, чем работа Гаузе над антибиотиками, известны его исследования по диссимметрии протоплазмы. Все три области исследований (диссимметрия протоплазмы, эволюционная теория, антибиотики) были для Гаузе логически связаны.

Гаузе воплощает несколько черт делающих его крайне интересным для настоящего исследования. Гаузе работал в условиях относительно изолированного тоталитарного государства, но в тоже время успешно публиковался на Западе, где очень рано получил известность. Тем не менее, именно его ранние экологические эксперименты вошли во все учебники и получили широкую известность, а его исследования антибиотиков и асимметрии протоплазмы известны в меньшей степени.

Поэтому количественные сравнительные исследования влияния Гаузе на отечественную и зарубежную науку представляют особенный интерес.

### IV. МЕТОДИКА ОТБОРА

Для первоначального отбора цифровых ресурсов были выбраны следующие критерии, расположенные по степени значимости в убывающем порядке:

- широта охвата по типам документов. При этом обязательным условием рассмотрения ресурсов должно быть наличие на них описаний научных статей. По этому признаку сразу же были исключены электронные каталоги библиотек;

- возможность выбора области поиска по метаданным, т.е. наличие выбора полей при формировании поискового запроса (например, название, аннотация, ключевые слова, годы публикаций и т.д.);

- достаточно глубокая временная ретроспектива содержащихся в ресурсе документов (в рамках проводимого исследования – вплоть до 1934 года);

- возможность выбора тематики публикаций, что позволяет получить более релевантные выборки документов;

- политематический характер ресурса для максимального охвата документов и публикаций.

Для верификации результатов поиска был использовано комбинированный поисковый запрос, состоящий из двух терминов «Gause» и «competitive exclusion principle», которые должны обязательно вместе содержаться в документе. При этом поиск должен проходить достаточно широко – по всем полям (название, аннотация, ключевые слова), а также при возможности и по тексту публикации.

### V. ЭМПИРИЧЕСКАЯ БАЗА ИССЛЕДОВАНИЯ

В соответствии с выбранными критериями и приведённой классификацией были отобраны следующие цифровые научные ресурсы.

SCOPUS (<http://scopus.com>) - политематическая реферативная база данных. Используемый поисковый запрос: Gause AND competitive exclusion principle (по заглавию, аннотации и ключевым словам, все года).

Всего было найдено 560 документов. При этом репрезентативность – с 1960 года.

Academic Search Complete (EBSCO, <https://www.ebscohost.com/academic/academic-search-complete>) является на данный момент самой полной мультидисциплинарной базой научных работ. Используемый поисковый запрос: Gause AND competitive exclusion principle (по всем полям и по тексту, диапазон годов 1934 – 2017). Всего было найдено 248 документов. При этом репрезентативность – с 1969 г.

Web of Science Core Collection (<http://webofknowledge.com>) - политематическая реферативная и наукометрическая база данных компании Thomson Reuters, охватывающая более 13 000 мировых научных журналов. Используемый поисковый запрос: Gause AND competitive exclusion principle (по теме, диапазон годов 1934 – 2017). Всего было найдено 9 документов. При этом репрезентативность – с 12005 г.

JSTOR (The Scholarly Journal Archive, <http://jstor.org>) – цифровые архивы, включающие научные материалы, опубликованные в более чем тысяче самых высококачественных академических журналах по гуманитарным, общественным и естественным дисциплинам, а также монографии и другие материалы. Используемый поисковый запрос: Gause AND competitive exclusion principle (полнотекстовый поиск, диапазон годов 1934 – 2017). Всего было найдено 1 722 документа. При этом репрезентативность – с 1934 г.

Один из самых крупнейших агрегаторов OAister (<http://oaister.worldcat.org>), который предоставляет возможность поиска по миллиону публикаций по всем отраслям знаний, которые доступны бесплатно в рамках «Open Archive Initiative». Основная их часть находится в университетских репозиториях «открытого доступа» и в растущем количестве журналов «открытого доступа». Используемый поисковый запрос: Gause AND competitive exclusion principle (поиск по ключевым словам, диапазон годов 1934 – 2017). Всего было найдено 6 документов. При этом репрезентативность – с 2002 г.

PLOS (Public Library of Science, <https://www.plos.org>) – ресурс, основанный в 2001 году и содержащий базу публикаций открытых журналов по медицине и биологии. Используемый поисковый запрос: Gause AND competitive exclusion principle (поиск по всем полям). Всего было найдено 68 документов. При этом репрезентативность – с 2008 г. Ресурс не обладает механизмом расширенного поиска. Ссылки выдаются общим списком без возможности применения дополнительных фильтров.

Google Scholar (<https://scholar.google.ru>) - поисковая система по полным текстам научных публикаций всех форматов и дисциплин. Проект работает с ноября 2004 года. Google Scholar включает данные из большинства рецензируемых электронных журналов крупнейших мировых научных издательств, а также из цифровых научных ресурсов свободного доступа (например, открытые репозитории и архивы). Используемый поисковый запрос: gause "competitive exclusion principle". Выявленная репрезентативность – с 1960 г.

Всего было найдено 1 930 документов. Ресурс не обладает механизмом расширенного поиска, но есть возможность фильтрации по релевантности и по годам публикации. Также можно исключать из результатов поиска цитирования. Экспериментально установлено, что для формирования поисковых запросов можно использовать элементы поискового метаязыка, реализованного в ИПС Google.

PubMed Central (<https://www.ncbi.nlm.nih.gov/pmc/>) - полнотекстовый архив свободного доступа статей из научных журналов по биомедицине и биологии. Используемый поисковый запрос: Gause AND competitive exclusion principle (поиск по всем полям). Всего было найдено 32 документа. Ресурс обладает механизмом расширенного поиска и применения дополнительных фильтров. Например, выбор диапазона годов публикаций.

OpenDOAR (<http://www.openoar.org>) – каталог открытых репозиториях (hosted by University of Nottingham, UK). Используемый поисковый запрос: Gause competitive exclusion principle. Всего было найдено 6 520 документов. Ресурс не обладает механизмом расширенного поиска. Отсутствует возможность использовать фильтр по годам публикаций. Ссылки выдаются общим списком без возможности применения дополнительных фильтров.

SpringerLink (<http://link.springer.com>) – электронная платформа издательства Springer, которая содержит более 8 млн. документов из более чем 2 700 научных журналов, более чем 4 000 сериальных изданий (с 1842 г.), а также более 100 000 книг (с 1842 г.) и другие документы. Используемый поисковый запрос: Gause competitive exclusion principle (поиск по всем словам, диапазон годов 1934 – 2017). Всего было найдено 453 документа. При этом репрезентативность – с 1958 г.

Для верификации всего доступного индексируемого цифрового информационного пространства аналогичные запросы были использованы в информационно-поисковых системах общего назначения Yandex (<http://yandex.ru>) и Google (<http://google.com>). При этом были выданы ссылки на 40 000 и 27 700 различных ресурсов соответственно.

Наглядное представление полученных результатов позволяет сравнить количественные показатели объёма данных, доступные через соответствующие информационные ресурсы (рис. 1).

После получения результатов по поисковым запросам был произведён анализ возможности выгрузки результатов, что является необходимым условием для дальнейшей их обработки. Оказалось, что возможностью выгрузки документов в том или ином виде обладают не все цифровые информационные ресурсы. Обобщённые результаты исследования возможностей рассмотренных цифровых ресурсов приведены в таблице 1.

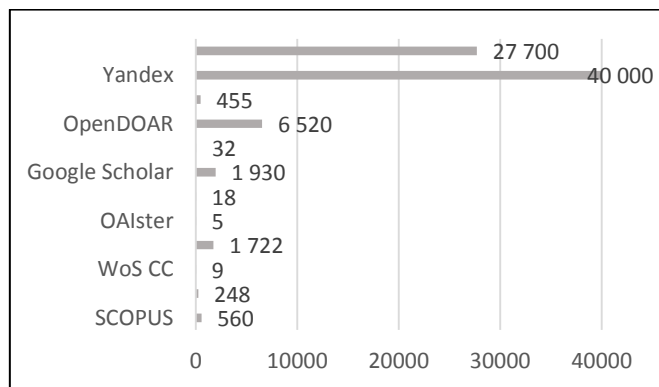


Рис. 1. Количество результатов выполнения поискового запроса «Gause AND competitive exclusion principle» в различных цифровых информационных ресурсах.

Таб. 1. Возможности различных цифровых ресурсов для количественной и качественной обработки полученных результатов.

| Источник                         | РАСПИРЕННЫЙ ПОИСК | НАЧАЛО РЕТРОСПЕКТИВЫ | ВОЗМОЖНОСТЬ ВЫГРУЗКИ РЕЗУЛЬТАТОВ       |
|----------------------------------|-------------------|----------------------|--|
| SCOPUS                           | Да                | 1969                 | Да (формат CSV)                        |
| Academic Search Complete (EBSCO) | Да                | 1969                 | Нет                                    |
| Web of Science Core Collection   | Да                | 2005                 | Да (формат CSV)                        |
| JSTOR                            | Да                | 1934                 | Да (RefWorks, EasyBib, RIS, Text file) |
| OAlster                          | Да                | 2002                 | Нет                                    |
| PLOS                             | Нет               | 2008                 | Нет                                    |
| Google Scholar                   | Нет               | 1960                 | Нет                                    |
| PubMed Central                   | Да                | 1972                 | Да (текст, XML)                        |
| OpenDOAR                         | Нет               | -                    | Нет                                    |
| SpringerLink                     | Да                | 1958                 | Да (формат CSV)                        |
| Yandex                           | Да                | -                    | Нет                                    |
| Google                           | Да                | -                    | Нет                                    |

## VI. ОЦЕНКА РОССИЙСКИХ ЦИФРОВЫХ ИНФОРМАЦИОННЫХ РЕСУРСОВ

Данное пилотное исследование было бы неполным без исследования как русскоязычного сегмента общедоступного массива информации (ИПС общего назначения и специализированные ИПС), так и российских цифровых информационных ресурсов. Для этих целей были использованы ИПС Google, Yandex, Академия Google, электронная библиотека «Киберленинка» и Научная электронная библиотека (НЭБ).

При этом с основным запросом «конкурентного исключения гаузе» использовались два термина – «принцип» и «закон». Это связано с тем, что в русскоязычном научном дискурсе встречаются для обозначения этого понятия термины «принцип» и «закон». Также в НЭБ был введён и прежде использованный запрос *gause* "competitive exclusion principle", по которому было получено 7 публикаций (все на русском языке), что говорит об ориентации этого ресурса на русскоязычные публикации.

Для ресурса Академия Google был использован запрос «"\* конкурентного исключения" гаузе» с использованием элементов поискового языка, являющегося универсальным инструментом во всех

поисковых сервисах Google. С учётом цитат было обнаружено 274 источника с репрезентативностью с 1984 года.

ИПС общего назначения Google и Yandex дали результат 12 200 и 6 000 соответственно с использованием запросов «"\* конкурентного исключения" гаузе» (Google) и «(закон от принцип) + "конкурентного исключения" + гаузе» (Yandex).

В поисковом сервисе Научной электронной библиотеки (НЭБ) был произведён поиск по фразе «конкурентного исключения гаузе», который дал всего 12 источников с репрезентативностью с 1991 года. При включении опций поиска по тексту и спискам цитируемой литературы выборка составила 183 источника с тем же началом репрезентативности.

Ещё меньше результатов выдала электронная библиотека Киберленинка – 80. При этом использовался достаточно общий запрос «конкурентного исключения гаузе», который позволил включить термины «принцип» и «закон». Достаточно скудные средства фильтрации позволили с использованием оператора @год методом простого перебора получить начало репрезентативности результатов, датируемое 2002 годом (этот оператор не позволяет задавать диапазон значений, что сделало бы процесс фильтрации более быстрым).

Обобщая результаты поиска по исследованным цифровым ресурсам, можно привести следующую таблицу, которая позволяет оценить их возможности для дальнейшего применения в науковедческих исследованиях (таб. 2).

Таб. 2. Результаты тематического поиска в русскоязычных цифровых информационных ресурсах

| Источник       | РАСПИРЕННЫЙ ПОИСК         | НАЧАЛО РЕТРОСПЕКТИВЫ | ВОЗМОЖНОСТЬ ВЫГРУЗКИ РЕЗУЛЬТАТОВ |
|----------------|---------------------------|----------------------|----------------------------------|
| Google Scholar | Частично, через операторы | 1984                 | Нет                              |
| Yandex         | Да                        | -                    | Нет                              |
| Google         | Да                        | -                    | Нет                              |
| НЭБ            | Да                        | 1991                 | Нет                              |
| Киберленинка   | Частично, через операторы | 2002                 | Нет                              |

## VII. ВЫВОДЫ

Первоначальная оценка возможностей цифровых информационных ресурсов позволяет выбрать приоритетные ресурсы для дальнейшего использования в целях исследования. Из рассмотрения по итогам анализа были исключены следующие ресурсы:

Academic Search Complete (EBSCO) – небольшое количество документов и отсутствие возможности выгрузки результатов поиска;

Web of Science Core Collection – небольшое количество документов;

OAlster – небольшое количество документов и отсутствие возможности выгрузки результатов поиска;

PubMed Central – небольшое количество документов;

PLOS – отсутствие расширенного поиска и применения дополнительных фильтров, небольшое количество документов и отсутствие возможности выгрузки результатов поиска;

PubMed Central – отсутствие расширенного поиска и

применения дополнительных фильтров, небольшое количество документов;

Отдельно необходимо отметить, что самые перспективные по количеству найденных публикаций ресурсы Google Scholar и OpenDOAR использовать для дальнейшего исследования не представляется возможным, так как у них ограничены инструменты расширенного поиска и дополнительной фильтрации результатов, а также отсутствуют возможности выгрузки результатов поиска.

По этим же причинам для дальнейшей качественной обработки результатов практически бесполезны ИПС Yandex и Google несмотря на впечатляющие количественные показатели. Также необходимо заметить, что в этих ИПС достаточно много как информационного шума, так и дублированной информации, что отличает их не в лучшую сторону по сравнению со специализированными базами научной информации, в которых обеспечивается уникальность каждого найденного документа.

В связи с этим наиболее адекватными для целей дальнейшего исследования можно считать такие цифровые ресурсы, как JSTOR, SCOPUS и SpringerLink. Помимо достаточно репрезентативной выборки документов эти ресурсы обладают достаточно удобными инструментами выгрузки полученных результатов для дальнейшей обработки математическими или статистическими методами.

Русскоязычные ресурсы можно использовать только для поиска русскоязычных источников. Эти ресурсы выдают значительно меньшее число результатов, что можно объяснить несколькими факторами:

- российские ресурсы появились и стали интенсивно пополняться информацией сравнительно недавно: НЭБ начала работать с русскоязычными документами с 2005 года, а Киберленинка была запущена в 2012 году;

- оба этих ресурса являются хранилищами, за наполнение которых отвечают издатели научных журналов и книг, размещающие информацию на своих условиях – для НЭБ (и Российского индекса научного цитирования, РИНЦ) это могут быть как полные тексты (журналы открытого доступа), так и платный контент (для подписных и коммерческих журналов и научной литературы); к тому издатели вовсе могут закрыть доступ к текстам. При этом обязательным условием для издателей является размещение в РИНЦ информации о выпусках журналов с 2005 года, а более ранние выпуски могут быть загружены по желанию издателя. Киберленинка преимущественно размещает статьи из журналов открытого доступа. В связи с перечисленными особенностями в этих ресурсах существуют лакуны, пропуски, отсутствуют тексты статей и других материалов, что, в общем, говорит о том, что они не отражают во всей полноте российские научные публикации;

- на российских ресурсах почти полностью отсутствуют научные публикации досоветского и советского периодов. А этот факт говорит, прежде всего, о том, что в России нет единой концепции сохранения национального научного наследия в цифровой форме и обеспечения доступа к нему;

- работа Г.Ф. Гаузе «The Struggle for Existence», в которой изложил принцип конкурентного исключения, была опубликована на английском в США в 1934 году. После этого в своей научной деятельности в СССР он данной тематики не касался, переключившись на другие исследования. Поэтому, скорее всего, до позднесоветского периода результаты исследований, изложенные в этой работе, не были широко известны советским учёным.

Дальнейшее исследование будет включать в себя обработку полученных массивов документов. Прежде всего необходимо сравнить результаты выгрузки и избавиться от дублирующих записей. Полученный таким образом обобщённый массив данных будет подвергнут дополнительным исследованиям.

#### БИБЛИОГРАФИЯ

- [1] Галл Я. М. Г. Ф. Гаузе — эколог и эволюционист. — СПб.: Альманах, 1997. — 160 с.
- [2] Галл Я.М. Георгий Францевич Гаузе. — СПб.: Нестор-История, 2012. — 236 с.
- [3] Гаузе Г.Ф. Лекарственные вещества микробов. — М., Л.: Издательство Академии Наук СССР, 1946. — 72с.
- [4] Гиляров А.М. Ранняя история русской экологии (неопубликованная рукопись).
- [5] Земсков А.И., Шрайберг Я.Л. Системы открытого доступа к информации: причины и история возникновения [Электронный текст] // Науч. и техн. б-ки. — 2008. №4. — URL: [http://ellib.gpntb.ru/subscribe/ntb/2008/4/ntb\\_4\\_2\\_2008.htm](http://ellib.gpntb.ru/subscribe/ntb/2008/4/ntb_4_2_2008.htm) (дата обращения: 24.02.2017)
- [6] Agosti, M., Benfante, L., Orio, N. IPSA: A Digital Archive of Herbs to Support Scientific Research // Digital Libraries: Technology and Management of Indigenous Knowledge for Global Access. ICADL 2003. Lecture Notes in Computer Science. 2003. — P. 2911. — doi:10.1007/978-3-540-24594-0\_24.
- [7] Borgman, C.L. The digital future is now: A call to action for the humanities // Digital Humanities Quarterly. 2009. Vol. 3 (4). — URL: <http://digitalhumanities.org/dhq/vol/3/4/000077/000077.html/000077.html>.
- [8] Borgman, C.L. Scholarship in the Digital Age: Information, Infrastructure, and the Internet. — MIT Press, 2007.
- [9] Burda, D., Teuteberg, F. Sustaining accessibility of information through digital preservation: A literature review // Journal of Information Science. 2013. Vol. 39 (4). — Pp. 442 – 458. — doi:10.1177/0165551513480107
- [10] Digital Libraries and Information Access: Research Perspectives. Edited by. G.G. Chowdhury and Schubert Foo. — Chicago, Illinois: Neal-Schuman/ALA, 2012.
- [11] Dublin Core. Российские цифровые библиотеки. — URL: [http://www.elbib.ru/index.phtml?env\\_page=methodology/metadata/md\\_review/md\\_descrip\\_general.html](http://www.elbib.ru/index.phtml?env_page=methodology/metadata/md_review/md_descrip_general.html)
- [12] Ferry, G. Scientific heritage: Science today, history tomorrow // Nature. 2013. № 493. — Pp. 19–21. — doi:10.1038/493019a.
- [13] Gause G.F. The Struggle for Existence. — Baltimore: Willkins, 1934.
- [14] Laakso, M. Green open access policies of scholarly journal publishers: A study of what, when, and where self-archiving is allowed // Scientometrics. 2014. Vol. 99 (2) — Pp. 475-494. — doi:10.1007/s11192-013-1205-3.
- [15] Lawal, I. Scholarly Communication: The Use and Non-Use of E-Print Archives for the Dissemination of Scientific Information // VCU Libraries Faculty and Staff Publications. 2002. №4. — doi:10.5062/F4057CWP.
- [16] Lourenço, M.C., Wilson, L. Scientific heritage: Reflections on its nature and new approaches to preservation, study and access // Studies in History and Philosophy of Science Part A. 2013. Vol. 44 (4). — Pp. 744-753. — doi:10.1016/j.shpsa.2013.07.011.
- [17] Nielsen, H.J., Hjørland, B. Curating research data: the potential roles of libraries and information professionals // Journal of Documentation. 2014. Vol. 70 (2). — Pp. 221-240. — doi:10.1108/JD-03-2013-0034.
- [18] OCLC: союз библиотек и удобство читателей // Российская государственная библиотека. — URL: <http://www.rsl.ru/ru/s7/s409/2013/20137642> (дата обращения: 14.02.2017).

- [19] OCLC. OAIster Contributors. URL: <http://www.oclc.org/oaister/contributors.en.html> (дата обращения: 12.02.2017).
- [20] Odlyzko, A. Competition and Cooperation: Libraries and Publishers in the Transition to Electronic Scholarly Journals // *Journal of Scholarly Publishing*. 1999. Vol. 30 (4). — Pp. 163-185. — doi:10.3138/JSP-030-04-163.
- [21] Pearce, N., Weller, M., Scanlon, E., Ashleigh, M. Digital Scholarship Considered: How New Technologies Could Transform Academic Work // *In Education*. 2010. Vol. 16 (1). — Pp. 33-44. — URL: <http://ineducation.couros.ca/index.php/ineducation/article/view/44>.
- [22] Scanlon, E. Scholarship in the digital age: Open educational resources, publication and public engagement // *Br J Educ Technol*. 2014. № 45. — Pp. 12–23. — doi:10.1111/bjet.12010.
- [23] Smith, I. Open access infrastructure. — Paris, UNESCO, 2015. — URL: <http://unesdoc.unesco.org/images/0023/002322/232204e.pdf>.
- [24] The Dublin Core Metadata for Simple Resource Discovery // Dublin Core Metadata Initiative. — URL: <http://dublincore.org/documents/>.
- [25] Weller, M. The Digital Scholar: How Technology Is Transforming Scholarly Practice. — London: Bloomsbury Academic, 2011. — doi:10.5040/9781849666275.

# Methods for Assessing the Possibility of Using Digital Information Resources in Scientific Heritage Studies (A Case Study of Georgy F. Gause's scientific ideas)

D.E. Prokudin, G.S. Levit

*Abstract* – In the information society, printed information will be increasingly digitized. To store the information, various scientific digital resources were created with access from the global information network. This makes it possible to use scientific information effectively for the purpose of promptly conducting the historical studies. The purpose of this case study is to explore the possibility of using digital information resources for application in the history of science. Our objective is to explore the ways of using digital databases in the history of science and society. We suggest selection methods suitable for historical studies and determine their effectiveness on the example of the influence of a Russian biologist Georgy F. Gause (1910-1986) and his evolutionary concepts. We demonstrate how various combinations of search criteria meet certain needs of science historians. We consider both global and Russian information resources. Finally, we provide general recommendations for searching and analyzing digital databases. The article reflects the further development of the case study, the first results of which were published in the materials of the International Scientific Conference Internet and Modern Society (IMS-2017) and the XIX All-Russian Scientific Conference "Scientific Service on the Internet".

*Key words:* Scientific information; digital information resources; methods of selecting information resources; criteria of selecting information resources; digital scientific heritage; Georgy F. Gause.