

ОГЛАВЛЕНИЕ

З. В. Апанович ИНФОРМАЦИЯ О РОССИЙСКИХ НАУЧНЫХ ОРГАНИЗАЦИЯХ В МЕЖДУНАРОДНЫХ И РУССКОЯЗЫЧНЫХ ИСТОЧНИКАХ ДАННЫХ	756–769
С. А. Власова, Н. Е. Каленов РАЗВИТИЕ ИНФОРМАЦИОННОЙ СИСТЕМЫ РЕГИСТРАЦИИ РЕЗУЛЬТАТОВ ИНТЕЛЛЕКТУАЛЬНОЙ ДЕЯТЕЛЬНОСТИ СОТРУДНИКОВ НАУЧНОГО УЧРЕЖДЕНИЯ	770–793
М. М. Горбунов-Посадов ПРОСПЕКТИВНАЯ БИБЛИОГРАФИЯ	794–807
Ф. О. Каспаринский ПРИНЦИПЫ МУЛЬТИТРЕЙДИНГА	808–869
А. С. Козицын, С. А. Афонин МЕТОД ПОИСКА ЭКСПЕРТОВ ПО ДАННЫМ НАУКОМЕТРИЧЕСКИХ СИСТЕМ	870–888
М. В. Михайлюк, Д. А. Кононов, Д. М. Логинов ТЕХНОЛОГИЯ СИТУАЦИОННОГО МОДЕЛИРОВАНИЯ В СИСТЕМАХ ВИРТУАЛЬНОГО ОКРУЖЕНИЯ	889–901
А. В. Никешин, В. З. Шнитман ОПЫТ ВЕРИФИКАЦИИ РЕАЛИЗАЦИЙ ПРОТОКОЛА TLS 1.3	902–922
А. А. Печников, Д. Е. Чебуков ИССЛЕДОВАНИЕ БИБЛИОГРАФИЧЕСКИХ ССЫЛОК В MATH-NET.RU С ИСПОЛЬЗОВАНИЕМ ГРАФА ЦИТИРОВАНИЯ ЖУРНАЛОВ	923–943
Ю. Е. Поляк ЗАМЕТКИ О ПЕРВЫХ ЭЛЕКТРОННЫХ БИБЛИОТЕКАХ	944–982

ОТ СОСТАВИТЕЛЕЙ

Настоящий тематический выпуск журнала «Электронные библиотеки» состоит из двух частей (первая часть – №5, вторая часть – №6) и включает статьи, подготовленные их авторами на основе материалов, представленных и доложенных на XXIII Всероссийской научной конференции «Научный сервис в сети Интернет». Конференция была проведена с 20 по 23 сентября 2021 года в режиме онлайн и традиционно была посвящена направлениям и тенденциям использования интернет-технологий в современных научных исследованиях. Основной целью конференции было предоставить возможность для обсуждения, апробации и обмена мнениями о наиболее значимых результатах в данной области деятельности, полученных ведущими российскими учеными за последнее время. Организатором конференции был Институт прикладной математики им. М.В. Келдыша Российской академии наук. Статьи, опубликованные в тематическом выпуске, в полной мере отражают самые современные результаты в области исследований, названной выше.

М.М. Горбунов-Посадов, А.М. Елизаров

УДК 004.021, 004.42

ИНФОРМАЦИЯ О РОССИЙСКИХ НАУЧНЫХ ОРГАНИЗАЦИЯХ В МЕЖДУНАРОДНЫХ И РУССКОЯЗЫЧНЫХ ИСТОЧНИКАХ ДАННЫХ

З. В. Апанович^[0000-0002-5767-284X]

*Институт систем информатики им. А.П. Ершова Сибирского отделения
Российской академии наук, пр. Академика Лаврентьева, 6, Новосибирск,
Новосибирская обл., 630090*

apanovich_09@mail.ru

Аннотация

Рассмотрены международные и русскоязычные источники данных, предоставляющие информацию о российских научных организациях. Продемонстрировано, что русскоязычные источники данных содержат больше информации о русскоязычных научных организациях, чем англоязычные, но эта информация остается недоступной для англоязычных источников данных. Также описаны эксперименты по сопоставлению и интеграции информации о российских научных организациях в международных и российских источниках данных. Рассмотрены такие источники данных, как GRID, русскоязычная и англоязычная Wikipedia, Wikidata и eLIBRARY.ru. Работа является промежуточным этапом на пути к созданию открытого и расширяемого графа знаний.

Ключевые слова: *разноязычные графы знаний, идентификация сущностей, научные организации, корректность*

ВВЕДЕНИЕ

Современные программные средства позволяют разрабатывать не только гетерогенные графы знаний, но и графы знаний, собирающие информацию о сущностях определенного класса [1]. Центральный банк Италии создал граф знаний обо всех итальянских компаниях [2], граф знаний ClaimsKG [3] извлекает утверждения из веб-страниц, занимающихся проверкой фактов, таких как Politifact (<https://www.politifact.com/>), и публикует их, используя модель RDF, а EventKG [4] извлекает информацию о месте и времени всех сущностей, относящихся к классу Event (событие).

Информация о научных организациях является важным атрибутом, позволяющим идентифицировать авторов научных публикаций [5, 6], а также анализировать географическое распределение публикаций и оценивать влияние на цитируемость публикаций, связанную с географическим фактором [7]. К сожалению, информация о национальных научных организациях, например, о российских, часто является неполной или искаженной в международных базах данных. Поэтому весьма актуальной является разработка открытого и расширяемого графа знаний, посвященного российским научным организациям.

Одной из больших открытых англоязычных баз данных научных организаций является база данных GRID [8] (Global Research Identifier Database, <https://www.grid.ac/>). Она содержит информацию о более чем ста тысячах научных организациях по всему миру и используется не только как самостоятельный источник данных: данные GRID интегрированы в граф знаний SN SciGraph, разрабатываемый издательством Springer (<https://www.springernature.com/gp/researchers/scigraph>). Благодаря использованию данных GRID, все информация о публикациях SN SciGraph имеет географическую привязку (координаты, город, район, область, страна). Вторым достоинством GRID является наличие ссылок (поле External links) на глобальные библиографические ресурсы, такие как ROR (Research Organization Registry, <https://ror.org>), Crossref (<https://www.crossref.org/>), и ISNI (International Standard Name Identifier, <https://isni.oclc.org/>). Кроме ссылок на глобальные библиографические ресурсы, в поле External links имеются адрес веб-сайта организации и, опционально, ссылка на страницу организации в англоязычной Wikipedia. Поле Alternate Labels содержит варианты названия организации на английском языке и национальном языке конкретной организации. Приводятся как варианты полных названий (поле Aliases), так и сокращенных (поле Acronyms). Наконец, имеется возможность выстраивания иерархии организаций при помощи таких свойств, как дочерняя организация (child institutes) и вышестоящая организация (parent institute).

Помимо этого, каждый институт в GRID имеет метаданные, такие как уникальный идентификатор организации в GRID, тип организации (например, правительственная) и время возникновения организации (поле Established).

Список российских научных организаций, представленных в GRID, весьма ощутимо пополнился в 2021 году. Если в начале 2021 года он содержал данные о 2019 российских организаций, то уже в начале октября 2021 года там имелось 2099 организаций. К сожалению, информация о российских научных организациях, хранящаяся в GRID, с одной стороны, неполна, с другой стороны, содержит явные неточности.

Например, в GRID имеется страница, посвященная Сибирскому отделению Российской академии наук (Siberian Branch of the Russian Academy of Sciences, <https://www.grid.ac/institutes/grid.415877.8>). В качестве русскоязычного названия этой организации указан «Институт космофизических исследований и аэронауки им. Ю.Г. Шафера Сибирского отделения Российской академии наук», а в качестве дочерних организаций (“child institutes”) наряду с несколькими институтами, формально относящимися к СО РАН, перечислены образовательные организации различного подчинения, не имеющие к СО РАН никакого отношения. Среди них, например, имеется Восточно-сибирский институт Министерства внутренних дел Российской Федерации (East-Siberian Institute of the Ministry of Internal Affairs of the Russian Federation (<https://www.grid.ac/institutes/grid.445063.0>), Сибирский юридический институт ФСКН России (<https://www.grid.ac/institutes/grid.445537.4>) и др. Также можно заметить, что база данных GRID содержит в своем списке организации, которые уже прекратили свое существование. Например, у Федерального агентства научных организаций (Federal Agency for Scientific Organizations, <https://www.grid.ac/institutes/grid.484124.f>) указана дата возникновения (1913 год), но нет информации, что эта организация упразднена в 1918 году.

Наблюдая подобные неточности, можно предположить, что данная база пополнялась на основании англоязычных источников данных и, возможно, были бы полезными сопоставление и интеграция этих данных с русскоязычными источниками. Поскольку в GRID уже имеется возможность указания страницы организации в англоязычной Wikipedia, возникает предположение, что русскоязычная Wikipedia может содержать больше информации о российских научных организациях, и эта информация может быть лучшего качества, чем англоязычные источники данных.

Поэтому первый эксперимент состоял в установлении соответствия между страницами GRID, а также англоязычной и русскоязычной версиями Wikipedia при

помощи Wikipedia_API. Первоначально были проверены имеющиеся в GRID ссылки на англоязычную версию Wikipedia. Из 2019 российских организаций, имеющих в GRID, только 412 имели ссылки на страницы в англоязычной Wikipedia, и только 398 страниц англоязычной Wikipedia имели межязыковые ссылки (interlanguage links) на страницы русскоязычной Wikipedia. Для выяснения того, имеются ли в англоязычной и русскоязычной Wikipedia другие страницы русскоязычных организаций, описанных в GRID, осуществлялся поиск по таким атрибутам, как URL веб-сайта организации, название организации и др. К сожалению, поиск в Wikipedia по адресу веб-сайта организации оказался не очень эффективным, и в конечном итоге пришлось осуществлять поиск по различным версиям названия организации. Поиск осуществлялся как по англоязычному, так и по русскоязычному названию в обеих языковых версиях Wikipedia. В результате этого поиска было дополнительно обнаружено 674 страницы в русскоязычной Wikipedia, соответствующие организациям из GRID, из которых 353 страницы были связаны межязыковыми ссылками с англоязычной Wikipedia. В общей сложности было обнаружено 835 страниц Wikipedia, соответствующих русскоязычным организациям, представленным в GRID. Таким образом, этот эксперимент показал, что в русскоязычной версии Wikipedia действительно имеется больше информации о российских научных организациях, чем в англоязычной, но эта информация остается недоступной для англоязычных баз данных из-за их ориентированности на англоязычную версию Wikipedia.

Основная проблема состояла в том, что в русскоязычной Wikipedia оказалось мало явных ссылок на веб-сайты организаций, а поиск по названиям организаций сильно затруднен тем, что в разных базах данных приводятся разные названия одних и тех же организаций. Существующий, очень несовершенный алгоритм сопоставления планируется доработать.

Другим русскоязычным источником данных, заслуживающим внимания, является самая большая база данных российских научных организаций eLIBRARY.ru [9], на сайте которой представлено 12420 российских организаций.

1. СРАВНЕНИЕ ИНФОРМАЦИИ В GRID И eLIBRARY.RU

При сравнении информации в GRID и eLIBRARY.ru бросается в глаза, что количество организаций, представленных на сайте eLIBRARY.ru, в пять раз превышает количество организаций, представленных в GRID. Это соотношение остается постоянным для разных субъектов РФ. Так, если eLIBRARY.ru выдает список из 264 организации, расположенных в Новосибирске, то GRID показывает только 75 таких организаций, причем некоторые из представленных организаций относятся к разряду несуществующих.

На основе сопоставления количества данных о российских организациях, имеющихся в eLIBRARY.ru и GRID, возникает вопрос, за счет чего формируется такое значительное различие. При внимательном рассмотрении было обнаружено, что eLIBRARY.ru действительно содержит очень большой список российских организаций, но далеко не все организации из имеющегося списка относятся к научным. В список организаций входят все федеральные министерства и организации, подчиненные этим министерствам, региональные администрации, а также коммерческие организации, банки, больницы, индивидуальные предприниматели, заводы, комбинаты и прочие. Например, в eLIBRARY.ru можно обнаружить такую организацию, как «Домостроительный комбинат № 7», за которой не числится ни одной публикации и не имеется никакой информации, кроме почтового и юридического адресов. В целом в списке организаций eLIBRARY.ru примерно одна треть (4505 организаций) вообще не имеет никаких публикаций, а еще примерно тысяча организаций имеет по одной публикации и столько же – по две публикации.

В этой базе данных также имеется немало организаций, прекративших свое существование. Например, у многих институтов СО РАН в качестве ведомства, которому подчиняется данная организация, указано Федеральное агентство научных организаций (ФАНО). Страничка ФАНО имеется и в базе GRID, но в русскоязычной Wikipedia имеется информация об упразднении этой организации 15 мая 2018 года. Эта же информация продублирована в наборе данных Wikidata (Federal Agency for Scientific Organizations, Q16711297).

Заметим, что алгоритм сопоставления данных из eLIBRARY.ru и GRID на основе названий организаций обнаружил эквиваленты только для половины российских организаций, указанных в GRID. Предполагается, что такой невысокий

процент связан, прежде всего, с реальным различием в информации об организациях. В GRID, как и в eLIBRARY.ru, обнаружались организации, прекратившие свое существование, а также и заведомо ошибочные данные об организациях. С другой стороны, алгоритм сопоставления данных нуждается в улучшении. Наконец, возникает естественное желание использовать в качестве посредника между русскоязычными и англоязычными источниками данных ресурс, имеющий доступ к разноязычным источникам данных.

2. СОПОСТАВЛЕНИЕ ДАННЫХ eLIBRARY.RU И WIKIDATA

Примером очень перспективного источника данных является набор данных Wikidata. Wikidata возникла в 2014 году [10] как источник структурированных данных для управления фактами в различных языковых версиях Wikipedia. Разработчики планируют сделать ее центральной платформой управления Wikipedia, интегрируя данные из всех языковых «глав» Wikipedia. Интеграция осуществляется присваиванием не зависящего от конкретной языковой версии идентификатора каждому объекту реального мира и объединения всех высказываний о заданном объекте реального мира из всех языковых версий Wikipedia. Например, на странице Wikidata, посвященной СО РАН (<https://www.wikidata.org/wiki/Q3032414>), имеются ссылки на описание этой сущности в шести языковых версиях Wikipedia. Помимо интеграции информации из различных языковых версий Wikipedia этот набор данных так же, как и GRID, поддерживает ссылки на глобальные библиографические ресурсы при помощи указания идентификаторов в таких источниках данных, как VIAF (VIAF ID, свойство P214), идентификатор ISNI (ISNI ID, свойство P213), идентификатор Библиотеки Конгресса США (Library of Congress authority ID, свойство P244), идентификатор GRID (GRID ID, свойство P2427), идентификатор ROR (ROR ID, свойство P6782), основной государственный регистрационный номер (ОГРН, свойство P7011), идентификатор организации в eLIBRARY.ru (eLIBRARY.ru organisation ID, свойство P2463).

Также, помимо официального названия организации, в наборе имеются короткое название на русском и английском языках, информация о географическом положении (страна, область), даты возникновения и исчезновения сущности. Несмотря на эти явные достоинства, в наборе данных имеются и недостатки. Напри-

мер, на странице, посвященной СО РАН, приводится тот же самый неточный список подчиненных институтов, что и в базе данных GRID. В нем есть образовательные организации других ведомств, но зато не хватает некоторых институтов, реально подчиняющихся СО РАН. Например, в списке институтов, подчиненных СО РАН, нет ИСИ СО РАН, зато ИСИ СО РАН представлен как институт РАН, и на его странице (<https://www.wikidata.org/wiki/Q4201722>) почему-то указано, что институт назван не в честь академика А.П. Ершова, известного своими работами по информатике, а в честь Александры Петровны Ершовой (Alexandra Petrovna Ershova, Q60830445) – российского театрального педагога.

Следует отметить, что весьма незначительное количество российских организаций, представленных в наборе данных Wikidata, имеет на своих страницах идентификаторы из eLIBRARY.ru. Так, запуская запрос SPARQL относительно научных организаций (wd:Q16519632), расположенных на территории России и имеющих в wikidata идентификатор eLIBRARY.ru, мы получили список из восьмидесяти шести таких организаций, причем указаны, в основном, образовательные организации. Новосибирский государственный университет, например, имеет идентификатор 214. Пример запроса, выдающего научные организации с сайта Wikidata с идентификатором eLIBRARY.ru, показан на рисунке 1.

В то же время, среди 274 организаций, относящихся в наборе данных Wikidata к институтам РАН (wd:Q4201890), только четыре организации имеют идентификатор eLIBRARY.ru. Это означает, что задача сопоставления данных в вышеуказанных источниках является весьма актуальной. Далее будет описан наш подход к решению этой проблемы.

```
select ?s ?s_label ?elib
where {
  ?s wdt:P31/wdt:P279+ wd:Q16519632;
  wdt:P17 wd:Q159;
  wdt:P2463 ?elib;
  rdfs:label ?s_label filter(lang(?s_label) = 'ru').
}
```

Рис. 1. Запрос, выдающий научные организации, имеющие идентификатор eLIBRARY.ru на сайте wikidata.

Заметим, что неполнота данных, хранящихся в Wikidata, не позволяет в настоящий момент получить всю достоверную информацию при помощи запроса SPARQL. Так, например, запрос SPARQL, требующий выдать все российские научные организации, выдает только те организации, у которых в явном виде указано, что они расположены в России, а если сделать требование российской принадлежности необязательным (OPTIONAL), начинают выдаваться лишние организации, у которых вообще не указана страна принадлежности. Поэтому приходится программно решать задачу сопоставления сущностей, описанных в разных источниках данных.

На вход программы интеграции данных подается таблица, каждая строка которой соответствует одной организации в eLIBRARY.ru. В качестве атрибутов, извлекаемых из eLIBRARY.ru, берутся такие атрибуты, как полное название организации на русском языке, название на английском языке, русская аббревиатура названия, английская аббревиатура названия, страна, регион, русское название города, английское название города, почтовый адрес на русском языке, почтовый адрес на английском языке, юридический адрес, ведомство, тип организации, факс, URL сайта.

В результате работы алгоритма интеграции создается расширенная таблица, которая, кроме данных из eLIBRARY.ru, содержит для каждой организации информацию, удалось ли обнаружить ее в Wikidata. В случае, если организация обнаружена, данные об организации дополняются информацией, извлеченной из Wikidata. В частности, добавляются Wikidata_имя, Wikidata_идентификатор, Wikidata_алиасы, Wikidata_год основания, VIAF_Id, eLIBRARY_Id, GRID_Id. В качестве возможного результата также может появиться сообщение «организация не обнаружена» и «данных для идентификации недостаточно».

Общая схема алгоритма интеграции имеет вид:

1. Предварительная обработка названий организаций, имеющихся в eLIBRARY.ru (перевод в нижний регистр, удаление стоп-слов «зао», «ооо», «им», циклическая замена некоторых слов при помощи словаря синонимов).

2. Поиск сущностей по одному из названий организации, указанных в eLIBRARY.ru (4 варианта названия, трансформированные названия, URL организации). В случае успешного поиска выдается файл JSON с краткой информацией о

найденном элементе, который позволяет извлечь дополнительную информацию о сущности: название, идентификатор, тип сущности, все доступные названия. В Wikidata эта информация содержится в поле «Также известный как».

3. Проверка, является ли сущность, найденная в Wikidata, эквивалентом организации из eLIBRARY.ru Для этого вычисляется коэффициент совпадения полных и трансформированных названий сущностей, выполняются сравнение URL организаций, проверка типа сущности, страны организации, местонахождения организации.

4. Пополнение исходной таблицы информацией из Wikidata

Сравнение названий организаций. При сравнении двух списков названий, извлеченных из eLIBRARY.ru и Wikidata, сначала осуществляется проверка на полное текстовое совпадение названий. В случае отрицательного результата проверки каждая строка-название разбивается на отдельные слова, они приводятся к нормальному виду при помощи модуля морфологического анализа. Для полученных строк вычисляется коэффициент совпадения по формуле:

$$Index = 2 \cdot (\text{количество совпадающих слов}) / (\text{суммарное количество слов}).$$

Вычисляется наилучший коэффициент совпадения по всем вариантам названий. Если этот коэффициент больше, чем 0,68 для некоторой пары названий, результат считается положительным, и осуществляется сравнение остальных атрибутов.

Сравнение URL организаций. При сравнении URL-адресов было замечено, что в базах данных он может быть записан по-разному. Например, в карточке организации *Администрация Архангельской области* в eLIBRARY.RU в качестве URL указан <http://www.dvinaland.ru/>, а в Wikidata – <https://dvinaland.ru/>. Таким образом, из-за различия конструкций “http” и “https” и наличия в одном из адресов конструкции “www” результат сравнения будет отрицательным. Во избежание таких ситуаций при сравнении URL указанные конструкции удаляются.

Проверка типа найденного объекта. Необходимо убедиться, что найденный объект является организацией. При помощи запроса SPARQL сгенерирован файл в формате CSV, содержащий названия всех подклассов класса Организация в онтологии Wikidata. Тип найденной сущности сравнивается с элементами этого файла. В случае отрицательного результата сравнения сущность отклоняется.

Проверка местонахождения организации. Если в качестве страны организации указана не Россия, поиск останавливается, в базу записывается, что организация находится не в России. Если Россия или другая страна не указаны на странице Wikidata, извлекается информация о местонахождении штаб-квартиры — в Wikidata чаще всего указывается город, также извлекается информация про административно-территориальное образование. Заметим, что в Wikidata вместо Москвы может быть указана, например, Московская область, что может привести к неправильному результату сравнения. Для решения этой проблемы используется информация об иерархии географических объектов. Был создан файл в формате JSON с распределением всех городов России по областям. Программа извлекает название города из Wikidata и, если оно не совпадает с положением, указанным в eLIBRARY.RU, то название города будет заменено на название области, в которой находится город. После этого географические положения сравниваются повторно. Если и в таком случае местоположения не совпадают, организация считается найденной неверно, и рассматривается следующая.

Таким образом, в настоящий момент организация считается верно идентифицированной в двух случаях:

А) Совпадают URL сайтов и некоторые пары вариантов названий.

Б) Сущность является организацией, названия организации в двух источниках совпадают, информация о сайтах не полна, организация находится в России. Для всех организаций, распознанных алгоритмом как идентичные, информация об этих организациях объединяется на основе расширения онтологии `schema.org`. На данный момент установлено соответствие между 3143 организациями Wikidata и eLibrary.ru. Получившийся экспериментальный источник данных будет в дальнейшем расширяться и интегрироваться с другими источниками данных, такими как GRID.

3. ЗАКЛЮЧЕНИЕ

Эксперименты с англоязычными и русскоязычными источниками данных показали, что русскоязычные источники информации содержат больше информации о русскоязычных научных организациях, чем англоязычные. К сожалению, эта информация остается в большинстве своем недоступной для англоязычных источ-

ников данных. Для решения этой проблемы разработан метод интеграции информации из разноязычных источников данных. Создана экспериментальная версия базы данных научных организаций, состоящая из 3143 научных организаций. Планируется превратить эту базу в открытый и расширяемый граф знаний. Авторы также считают, что для поддержания полноты и корректности информации о научных организациях каждая научная организация должна поддерживать на международных платформах свою страницу с указанием всех идентификаторов организации.

СПИСОК ЛИТЕРАТУРЫ

1. *Апанович З.В.* Эволюция понятия и жизненного цикла графов знаний // Системная информатика. 2020. № 16. С. 57–74.
 2. *Atzeni P., Bellomarini L., Iezzi M., Sallinger E., Vlad A.* Weaving Enterprise Knowledge Graphs: The Case of Company Ownership Graphs. URL: https://openproceedings.org/2020/conf/edbt/paper_334.pdf
 3. *Tchechmedjiev A., Fafalios P., Boland K., Gasquet M., Zloch M., Zapilko B., Dietze S., Todorov K.* ClaimsKG: A knowledge graph of fact-checked claims. In International Semantic Web Conference, Springer, 2019. P. 309–324.
 4. *Gottschalk S., Demidova E.* EventKG: A Multilingual Event-Centric Temporal Knowledge Graph. URL: <https://arxiv.org/pdf/1804.04526.pdf>
 5. *Apanovich Z.* Matching of authors and publications in multilingual bibliographic knowledge bases // CEUR Workshop Proceedings. SSI 2019. Proceedings of the 21st Conference on Scientific Services and Internet. 2020. P. 26–37.
 6. *Hajra A., Radevski V., Tochtermann K.* Author profile Enrichment for Cross-linking Digital Libraries // Research and Advanced Technology for Digital Libraries Springer International Publishing. Lecture Notes in Computer Science. 2015. Vol. 9316. P. 124–136.
 7. *Mannocci A., Osborne F., Motta E.* Geographical trends in academic conferences: An analysis of authors' affiliations // Data Science. 2019. 2(1) P. 181–203.
 8. Global Research Identifier Database. URL: <https://www.grid.ac/>
 9. Научная электронная библиотека eLIBRARY.ru.
URL: <https://www.elibrary.ru/>
-

10. *Ismayilov A., Kontokostas D., Auer S., Lehmann J., Hellmann S.* Wikidata through the Eyes of DBpedia.

URL: <http://www.semantic-web-journal.net/system/files/swj1462.pdf>

11. *Lenzerini M.* Data Integration: A Theoretical Perspective Conference: Proceedings of the Twenty-first ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems. 2002. P. 233–246.

INFORMATION ABOUT RUSSIAN RESEARCH ORGANIZATIONS IN MULTILINGUAL DATA SOURCES

Z. V. Apanovich^[0000-0002-5767-284X]

A.P. Ershov Institute of Informatics Systems, Siberian Branch of the Russian Academy of Sciences, Novosibirsk State University, Novosibirsk

apanovich_09@mail.ru

Abstract

International and Russian-language data sources that provide information about Russian research-related organizations are considered. It is demonstrated that Russian-language data sources contain more information about Russian research-related organizations than most international data sources, but this information remains unavailable for English-language data sources. Experiments on comparison and integration of information about Russian research organizations in international and Russian data sources are outlined. Data sources such as GRID, Russian and English chapters of Wikipedia, Wikidata and eLIBRARY.ru are considered. The work is an intermediate step towards the creation of an open and extensible knowledge graph.

Keywords: *multi-lingual knowledge graphs, identity resolution, research-related organizations, correctness*

REFERENCES

1. *Apanovich Z.V.* Evolyuciya ponyatiya i zhiznennogo cikla grafov znaniy // *Sistemnaya informatika*. 2020. № 16. S. 57–74.

2. *Atzeni P., Bellomarini L., Iezzi M., Sallinger E., Vlad A.* Weaving Enterprise Knowledge Graphs: The Case of Company Ownership Graphs. URL: https://openproceedings.org/2020/conf/edbt/paper_334.pdf

3. *Tchechmedjiev A., Fafalios P., Boland K., Gasquet M., Zloch M., Zapilko B., Dietze S., Todorov K.* ClaimsKG: A knowledge graph of fact-checked claims. In International Semantic Web Conference, Springer, 2019. P. 309–324.

4. *Gottschalk S., Demidova E.* EventKG: A Multilingual Event-Centric Temporal Knowledge Graph. URL: <https://arxiv.org/pdf/1804.04526.pdf>

5. *Apanovich Z.* Matching of authors and publications in multilingual bibliographic knowledge bases // CEUR Workshop Proceedings. SSI 2019 – Proceedings of the 21st Conference on Scientific Services and Internet. 2020. P. 26–37.

6. *Hajra A., Radevski V., Tochtermann K.* Author profile Enrichment for Cross-linking Digital Libraries // Research and Advanced Technology for Digital Libraries Springer International Publishing. Lecture Notes in Computer Science. 2015. Vol. 9316. P. 124–136.

7. *Mannocci A., Osborne F., Motta E.* Geographical trends in academic conferences: An analysis of authors' affiliations // Data Science. 2019. 2(1) P. 181–203.

8. Global Research Identifier Database. URL: <https://www.grid.ac/>

9. eLIBRARY.ru. URL: <https://www.elibrary.ru/>

10. *Ismayilov A., Kontokostas D., Auer S., Lehmann J., Hellmann S.* Wikidata through the Eyes of DBpedia.

URL: <http://www.semantic-web-journal.net/system/files/swj1462.pdf>

11. *Lenzerini M.* Data Integration: A Theoretical Perspective Conference: Proceedings of the Twenty-first ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems. 2002. P. 233–246.

СВЕДЕНИЯ ОБ АВТОРЕ



АПАНОВИЧ Зинаида Владимировна – старший научный сотрудник Института систем информатики СО РАН, доцент Новосибирского государственного университета. Сфера научных интересов – визуализация информации, визуализация графов, Semantic Web.

Zinaida Vladimirovna APANOVICH – senior researcher of the Institute of Informatics Systems of SB RAS, Associate Professor of Novosibirsk State University. Research interests include information visualization, graph visualization, Semantic Web.

email: apanovich@iis.nsk.su

ORCID 0000-0002-5767-284X

Материал поступил в редакцию 25 октября 2021 года

УДК 013, 004.65

РАЗВИТИЕ ИНФОРМАЦИОННОЙ СИСТЕМЫ РЕГИСТРАЦИИ РЕЗУЛЬТАТОВ ИНТЕЛЛЕКТУАЛЬНОЙ ДЕЯТЕЛЬНОСТИ СОТРУДНИКОВ НАУЧНОГО УЧРЕЖДЕНИЯ

С. А. Власова¹ [0000-0003-1533-5850], **Н. Е. Каленов**² [0000-0001-5269-0988]

¹⁻²Межведомственный суперкомпьютерный центр (МСЦ) РАН – филиал ФГУ ФНЦ Научно-исследовательский институт системных исследований (НИИСИ) РАН

¹vlas.svetlana2013@yandex.ru, ²nekalenov@mail.ru

Аннотация

Описана разработанная авторами веб-система, реализующая сервисы, связанные с формированием и предоставлением многоаспектной информации о результатах научной деятельности (публикациях, авторских свидетельствах и докладах на научных мероприятиях) сотрудников организации или группы организаций. Система ориентирована как на конечного пользователя, заинтересованного в получении конкретных данных, так и на административный персонал, формирующий отчетные материалы для вышестоящей организации. Информационная база системы содержит связанные данные о следующих классах объектов: персоны (авторы), организации и их подразделения; публикации на аналитическом, монографическом и сводном уровнях; авторские свидетельства; научные мероприятия (конференции, симпозиумы, семинары); доклады. В состав системы входят два модуля – административный, предназначенный для ввода и редактирования данных, и пользовательский, который представляет собой специальный поисковый аппарат, осуществляющий поиск информации, ее визуализацию, навигацию по связанным ресурсам и экспорт данных. Отличительной особенностью системы является введенное понятие «эквивалентных» объектов. Эквивалентными считаются объекты, представленные в системе различными метаданными, но относящимися к одной физической сущности. Такими объектами являются «персоны», соответствующие одному автору с различными написаниями фамилии в библиографических описаниях публикаций; организации, имеющие различные

варианты названий; статьи, опубликованные без изменений на различных языках. В соответствии с современными требованиями к отчетности по публикациям в системе отражаются источники финансирования научных исследований, а также аффилиации каждого автора, указанные в статьях.

Ключевые слова: *научные труды, научная деятельность, автоматизированная система, база данных, отчеты, сетевые технологии.*

ВВЕДЕНИЕ

Эффективность научных исследований, проводимых бюджетными научными организациями, базируется на оценках результатов интеллектуальной деятельности их сотрудников, отражаемой, в первую очередь, в научных публикациях и докладах на научных конференциях. Последние организационные решения в этой области предусматривают не только количественную, но и качественную оценку публикаций и докладов. В этой связи для каждой организации всё большую актуальность приобретают вопросы создания инструментария, позволяющего в автоматизированном режиме регистрировать результаты интеллектуальной деятельности сотрудников и оперативно формировать необходимые отчетные данные.

Учету публикаций сотрудников научных организаций уделялось значительное внимание, начиная с 1930-х годов, – все библиотеки академических институтов обязаны были вести картотеки трудов сотрудников. Но это была деятельность, в первую очередь, направленная на помощь читателям, которым периодически необходимо было предоставлять списки своих работ для прохождения аттестации, получения новой должности или звания и т. п. С появлением в библиотеках вычислительной техники картотеки трудов сотрудников стали заменяться базами данных, с которыми работали сотрудники библиотек. Развитие интернет позволило перейти к сетевым технологиям создания и поддержки баз данных публикаций сотрудников и существенно расширить область их применения [1–7]. Однако большинство поддерживаемых систем отражает публикации в «традиционном» виде – на основе стандартного библиографического описания. Одной из проблем, возникающих при этом, является различное написание фамилий авторов публи-

каций, в первую очередь транслитерируемых с кириллицы на латиницу. Для получения полного набора публикаций того или иного автора требуется формулировать запросы, содержащие все возможные варианты написания фамилий. Эта проблема характерна не только для локальных систем учета публикаций, но и для крупнейших систем мирового уровня. Так, фамилия Королёв в русском варианте иногда пишется как Королёв, а иногда – как Королев, что допустимо российскими правилами написания подобных фамилий (см. фрагмент результатов поиска в интернет информации о Сергее Королеве, представленный на рис. 1).

В публикациях, отраженных в зарубежных базах данных, фамилия Королёв транслитерируется в большинстве случаев как Korolev, иногда как Korolyov, иногда как Koroljov. Иллюстрация этого приведена на рис. 2: на запрос к базе данных Web of Science Core Collection (WoS CC) по автору «Korolev A.» выдается 1128 публикаций, «Korolyov A.» 107 публикаций, «Koroljov A.» – 5 публикаций.

Следует отметить, что возможность «интеграции» различных написаний фамилий одного автора реализуется в системе ORCID (Open Researcher and Contributor ID) [8]. Однако тестовый поиск показывает, что, по крайней мере, для ряда русских фамилий поисковая система ORCID работает неудовлетворительно. Так, при поиске по фамилии Королев система выдает 36 записей (рис. 3), а при поиске по фамилии Королёв – 9 записей (рис. 4).

[w Королёв, Сергей Павлович – Википедия](https://ru.wikipedia.org/wiki/Королёв,_Сергей_Павлович)
https://ru.wikipedia.org/wiki/Королёв,_Сергей_Павлович

< Обзор **Биография** История болезни и смерть Награды и зва >

 Сергей Королёв родился поздним вечером 30 декабря 1906 (12 января 1907) года в городе Житомире в семье учителя русской словесности Павла Яковлевича Королёва (1877–1929), родом из Могилёва, и дочери нежинского купца – Марии Николаевны Москаленко (Баланиной) (1888–1980) . Крестили его в Софийской церкви.
28 июня 1908 года семья Королёвых переехала в Киев, где отец получил место преподавателя рус...

Wikipedia · Текст по лицензии CC-BY-SA

Изображение Королев Сергей
<bing.com/images>



Посмотреть все изображения >

24 Сергей Королев - биография, личная жизнь, конструктор ...
<https://24smi.org/celebrity/3744-sergei-korolev.html>

< **Биография** Интересные Факты из Биогр... Первые шаги отечественног >

 Сергей Павлович Королев — выдающийся советский конструктор и ученый XX века, академик АН СССР, основоположник космонавтики, создатель программ и крупнейший специалист в области ракетостроения и кораблестроения. Сергей Королев родился 12 января 1907 года (по старому стилю 31 декабря 1906) в Житомире. Его отец был учителем, из разночинцев. После распада семьи мальчик...

Дополнительно: 24smi.org
4/5 ★★★★★ (69) Опубликовано: 14.01.2017

Королев Сергей Павлович 12 января 1907 - 14 января 1966 ...
<https://histrf.ru/lichnosti/biografii/p/koroliev-sierghiei-pavlovich>

12 января 1907 - 14 января 1966. Сергей Павлович Королёв (1907-1966) — выдающийся конструктор и ученый, работавший в области ракетной и ракетно-космической техники. Дважды Герой Социалистического Труда, лауреат Ленинской премии, академик Академии наук СССР, он является создателем ...

Рис. 1. Пример различного написания фамилии Королёв (Королев)



Рис. 2. Различная транслитерация фамилии Королев в WoS CC

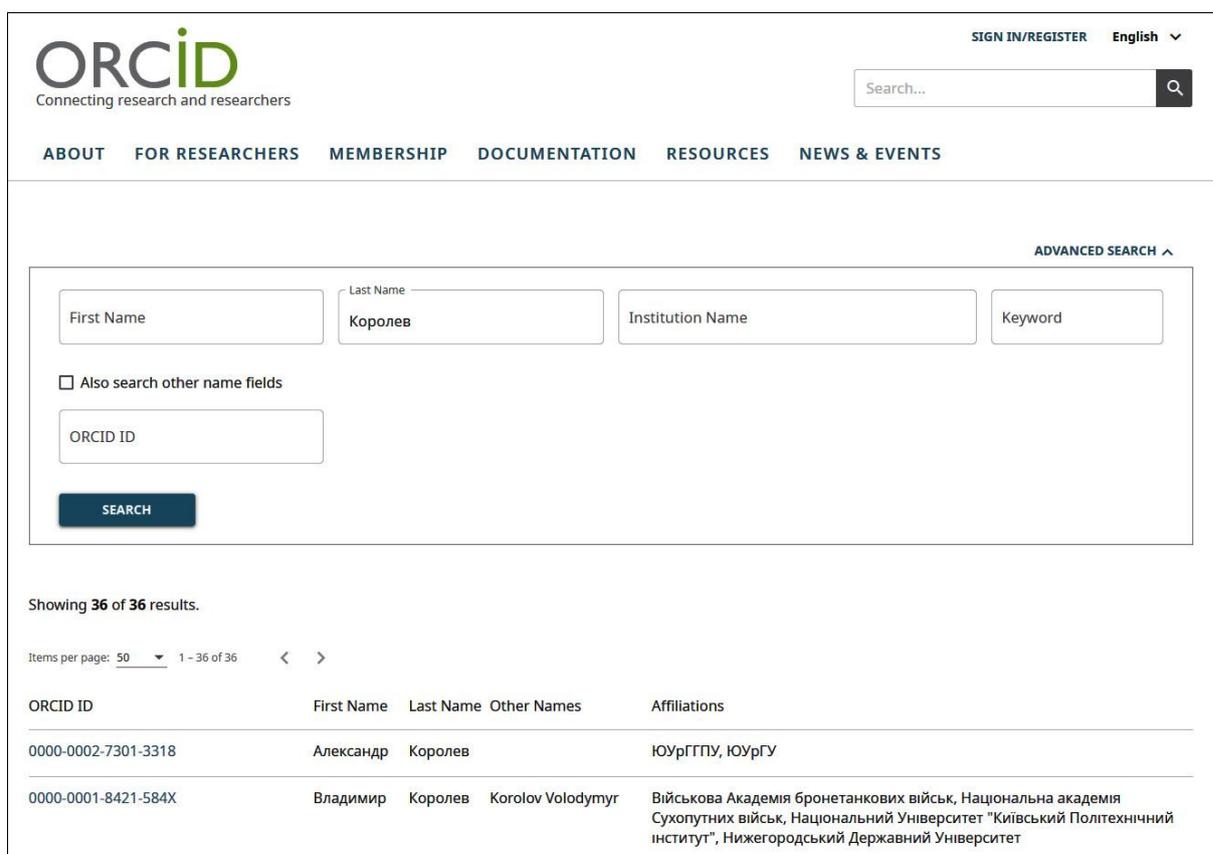


Рис. 3. Результат поиска по фамилии Королев в системе ORCID

The screenshot shows the ORCID website interface. At the top left is the ORCID logo with the tagline 'Мы связываем исследователей с их работами'. On the top right, there are links for 'ВОЙТИ/ЗАРЕГИСТРИРОВАТЬСЯ' and 'Русский'. A search bar is located in the top right corner. Below the header, there is a navigation menu with items: 'О', 'ДЛЯ НАУЧНЫХ РАБОТНИКОВ', 'ЧЛЕНСТВО В ORCID', 'ДОКУМЕНТЫ', 'РЕСУРСЫ', and 'НОВОСТИ И СОБЫТИЯ'. The main content area features a search form with fields for 'Имя', 'Фамилия' (filled with 'королёв'), 'Название учреждения', and 'Ключевое слово'. There is a checkbox for 'Также искать в других полях имен' and an 'ORCID ID' field. A 'поиск' button is at the bottom of the form. Below the search form, it says 'Показываются 9 из 9 результаты.' and 'Элементов на странице: 50' with pagination controls. The results are displayed in a table with columns: 'ORCID ID', 'Имя', 'Фамилия', 'Другие имена', and 'Связанные организации'.

ORCID ID	Имя	Фамилия	Другие имена	Связанные организации
0000-0003-4845-6042	Александр	Королёв		Omsk State Agrarian University, Omsk economics Institute, Semipalatinsker Pedagogische Hochschule
0000-0001-5461-0128	Алексей	Королёв		North Caucasian Institute of Mining and Metallurgical
0000-0001-9126-309X	Максим Геннадьевич	Королёв	Королёв МГ	Vitebsk State Medical University

Рис. 4. Результат поиска по фамилии Королёв в системе ORCID

Другой проблемой, связанной с анализом публикационной активности сотрудников той или иной организации, является переименование организации или изменение ее статуса, в частности, при объединении с другими организациями, что характерно для текущего момента реорганизации российской научной инфраструктуры. В международном масштабе эту проблему пытаются решить в рамках Research Organization Registry Community [9].

Одной из функций автоматизированной системы, регистрирующей в масштабах организации результаты интеллектуальной деятельности научных сотрудников, должно являться формирование отчетных данных, соответствующих требованиям Министерства науки и высшего образования. Согласно последним регламентирующим документам, данные о публикационной активности организации должны учитывать аффилиацию каждого автора, указанную в публикации, и источник финансирования исследований, которым посвящена та или иная статья.

В отчетах также должны указываться доклады на научных мероприятиях с указанием статуса мероприятия и доклада. Для оценки персонального вклада того или иного сотрудника в научную деятельность во многих организациях при формировании внутренних отчетов требуется указание, кто именно из соавторов доклада выступал на научном мероприятии.

Анализ современных публикаций и систем учета трудов научных сотрудников, представленных в интернет [10–12], показал, что ни одна из них не решает указанные выше проблемы, в том числе, и наиболее распространенная система ИСТИНА [13], знакомство с которой, в определенной степени, стимулировало разработку описываемой нами системы, которая решает практически те же задачи, но имеет существенно более развитый поисковый интерфейс.

Ниже представлено описание разработанной авторами системы учета результатов интеллектуальной деятельности (СУРИД), в которой реализованы перечисленные выше функции. Представленная система является результатом развития работ, ранее проведенных авторами в этом направлении [14, 15].

Первая версия системы регистрации публикаций сотрудников академических институтов была разработана и внедрена несколько лет назад в централизованной библиотечной системе БЕН РАН, где продолжает успешно эксплуатироваться вплоть до настоящего времени (<http://www.benran.ru/publben/>). Следующая версия системы разработана в МСЦ РАН в 2020 г., она была представлена в докладе на XXII Всероссийской научной конференции «Научный сервис в сети Интернет» и в опубликованных статьях [16, 17]. В этом году система получила свое дальнейшее развитие в связи с указанными выше задачами предоставления отчетных данных организации о результатах интеллектуальной деятельности научных сотрудников.

Новая версия системы обеспечивает регистрацию сведений о публикациях сотрудников, полученных ими авторских свидетельствах, докладах, сделанных ими на научных конференциях, симпозиумах, семинарах. Система предусматривает отражение данных, необходимых для формирования различных внутренних и внешних отчетов организации, а также позволяет решать перечисленные выше проблемы, связанные с неоднозначностью представления фамилий авторов публикаций и наименований организаций.

1. СТРУКТУРА СИСТЕМЫ

СУРИД обеспечивает создание и поддержку следующих взаимосвязанных объектов:

- публикации на аналитическом и монографическом уровнях;
- источники (издания на сводном уровне, в которых опубликованы статьи);
- доклады;
- научные мероприятия, на которых сделаны доклады;
- персоны (авторы публикаций и докладов);
- организации и их подразделения.

В новой версии системы по сравнению с предыдущей [16, 17] расширены профили метаданных объектов следующих классов.

Класс «Персона» – добавлены идентификаторы персоны в системах ORCID, РИНЦ, Scopus, WoS.

Класс «Источник» – добавлены номера ISSN, ISBN.

Класс «Публикация» – добавлены поля: номер государственного задания; информация о грантах, поддержавших исследования, приведенная в публикации.

Класс «Мероприятие» – добавлено поле статус мероприятия (российское, международное, региональное, локальное).

Класс «Доклад» – добавлены поля: язык (русский, английский); ссылка на презентацию доклада; ссылка на видеозапись выступления.

В новой версии системы вместо связи между объектами «Персона – организация», «Публикация – персона», «Доклад – персона» реализованы следующие связи:

Публикация – персона – организация (указание на аффилиацию автора);

Доклад – персона – подразделение организации;

Доклад – персона – докладчик;

Доклад – персона – содокладчик.

2. РЕГИСТРАЦИЯ ОБЪЕКТОВ В СИСТЕМЕ

Изменения в структуре метаданных и связей объектов обусловили необходимость переработки пользовательского интерфейса процесса регистрации публикаций и докладов. Рассмотрим подробно данные процессы.

Ввод данных публикации начинается с ввода ее авторов в том порядке, который представлен в публикации. Сначала нужный автор ищется в базе данных системы путем ввода в поисковую строку начального фрагмента фамилии. Система выдаст список найденных персон (фамилии являются активными ссылками), а также ссылку «Новая персона». Если автор публикации присутствует в данном списке, то нужно активизировать ссылку на его фамилии. Система покажет фамилию персоны и относящиеся к ней организации – названия организаций, которые были зарегистрированы ранее при вводе публикаций данного автора (рис. 5).

Ввод новой публикации

[Атаева О.М.](#)
Федеральный исследовательский центр Информатика и управление РАН

[Каленов Н.Е.](#)
 Библиотека по естественным наукам РАН
 Научно-исследовательский институт системных исследований | Межведомственный суперкомпьютерный центр Российской академии наук | Отдел информационных ресурсов и систем

[Ещё организация для данной персоны](#)

Рис. 5. Ввод информации об авторе и его аффилиации в новую публикацию.

Если нужной организации нет в списке, то ее можно добавить, перейдя по ссылке «Ещё организация для данной персоны». После перехода по данной ссылке система запросит фрагмент названия организации для ее выбора из введенных ранее (организацию можно найти по фрагментам ее полного и краткого названий). Оператор выбирает организацию из предложенного списка или вводит новую, которая автоматически добавляется к списку организаций для вводимого автора.

В случае отсутствия нужной персоны в системе, ее нужно ввести, перейдя по ссылке «Новая персона».

Для «привязки» персоны к публикации нужно выбрать (отметить «галочкой») организацию (или несколько организаций), которая указана в качестве аффилиации данного автора во вводимой публикации, и нажать на кнопку «Ввод автора в публикацию».

Авторов, введенных в публикацию, система показывает в виде списка фамилий в порядке их ввода в публикацию с указанием выбранных организаций. Фамилии являются активными ссылками, переход по которым позволяет при необходимости удалить или заменить автора. После окончания ввода всех авторов новой публикации система предоставит форму для ввода ее метаданных (рис. 6). В зависимости от вида публикации (статья, монография, авторское свидетельство, др.) осуществляется автоматический контроль обязательности ввода тех или иных полей данных. Обязательными для заполнения при вводе любого вида публикации являются название и год издания. При вводе статьи обязательным является указание на источник (журнал, сборник). Интерфейс ввода источника аналогичен описанному выше интерфейсу ввода организации, связанной с автором. После окончания ввода всех необходимых метаданных публикация будет зарегистрирована в системе.

К докладу нужно привязать мероприятие, на котором он был сделан. По фрагментам названия мероприятия определяется его наличие в системе. В случае его отсутствия предоставляется форма для ввода метаданных мероприятия: название; вид (конференция, семинар, симпозиум, совещание, конгресс); статус мероприятия (российское, международное, региональное, локальное); место проведения: страна, город; сроки проведения: дата начала – дата окончания; адрес сайта с информацией о мероприятии; ссылка на опубликованные материалы мероприятия (рис. 8).

Регистрация нового доклада так же, как и ввод новой публикации, начинается с ввода его авторов. Поиск авторов, выбор для них организаций и их привязка к докладу происходят аналогично вводу авторов публикации. Для каждого автора доклада указывается статус: докладчик или содокладчик. Затем система предоставляет форму для ввода метаданных доклада: название доклада; вид доклада (пленарный, секционный, стендовый, приглашенный); язык (русский, английский); адрес презентации доклада; адрес видеозаписи выступления. На рис. 7

представлен пример ввода метаданных доклада (фамилия докладчика выделена жирным шрифтом).

Атаева О.М. <i>Федеральный исследовательский центр Информатика и управление РАН (ФИЦ ИУ РАН)</i>	
Каленов Н.Е. <i>Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)</i>	
Серебряков В.А. <i>Федеральный исследовательский центр Информатика и управление РАН (ФИЦ ИУ РАН)</i>	
Добавить автора	
Название статьи, библиографическое описание монографии, авторского свидетельства	Онтологический подход к описанию единого цифрового пространства научных знаний
Вид публикации	статья в журнале
Источник	Электронные библиотеки Удалить источник
Год издания	2021
Том	24
Номер	1
Страницы	3-19
Язык публикации	русский
Идентификаторы во внешних базах данных (DOI, WOS, Scopus)	DOI: 10.26907/1562-5419-2021-24-1-3-19
Номер государственного задания	
Гранты (РФФИ, РГНФ)	РФФИ 20-07-00324, 18-00-00297, 18-00-00372
Адрес полного текста	publ/HH2214.pdf

Рис. 6. Регистрация публикации

Власова С.А. <i>Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)</i>	
Каленов Н.Е. <i>Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)</i>	
Добавить автора	
Название доклада	Информационная система «Научные труды сотрудников академических учреждений»
Вид доклада	пленарный
Язык	русский
Мероприятие	XXII Всероссийская научная конференция "Научный сервис в сети Интернет"
Презентация	/bd/pr/2020/42.pptx
Видеозапись выступления	/bd/mp4/2020/42.mp4
Дополнительная информация	

Рис. 7. Регистрация доклада

Название мероприятия	<input type="text" value="XXII Всероссийская научная конференция " в="" интернет"="" научный="" сервис="" сети=""/>
Вид мероприятия	<input type="text" value="конференция"/>
Статус мероприятия	<input type="text" value="российское"/>
Место проведения (страна)	<input type="text" value="Россия"/>
Место проведения (город)	<input type="text" value="онлайн"/>
Даты мероприятия	год <input type="text" value="2020"/> с <input type="text" value="21"/> . <input type="text" value="09"/> по <input type="text" value="25"/> . <input type="text" value="09"/> (число.месяц)
Ссылка на сайт мероприятия	<input type="text" value="http://agora.guru.ru/display.php?conf=abrau2020"/>
Ссылка на материалы	<input type="text" value="https://keldysh.ru/abrau/2020/proc.pdf"/>
Дополнительная информация	<input type="text"/>

Рис. 8. Регистрация мероприятия

3. ПОЛЬЗОВАТЕЛЬСКИЙ МОДУЛЬ СИСТЕМЫ

Пользовательский блок системы (<http://dirsmc.ru/bd/>) представляет собой поисковый аппарат, обрабатывающий запросы различной сложности. Запросы могут включать элементы всех атрибутов профилей метаданных объектов, объединенные операторами булевой логики «И», «ИЛИ», «И НЕ». Логика составления и выполнения системой запросов по различным поисковым полям подробно описана в [16]. Здесь мы остановимся на интерфейсе предоставления пользователю найденной информации и навигации по связанным ресурсам.

Система предоставляет пользователю возможность указать, о каких классах объектов он хочет получить информацию в ответ на свой запрос непосредственно в результирующей выдаче. Это могут быть данные о публикациях, источниках (журналах, сборниках), докладах, мероприятиях, персонах или организациях. Поисковый интерфейс позволяет обрабатывать запросы типа «найти журналы и сборники, в которых в период 2018–2020 гг. были опубликованы статьи сотрудников отдела информационных ресурсов и систем (ОИРС) МСЦ, поддержанные грантами РФФИ». На этот запрос (рис. 9) выдаются 7 наименований журналов и сборников, каждое из которых является активной ссылкой, при переходе по которой выдается список всех публикаций из данного журнала (сборника), имеющих в системе.

БД научных трудов сотрудников МСЦ РАН

Система позволяет получать и просматривать списки публикаций, журналов, где они опубликованы, авторов (персоны), связанных с ними организаций, докладов и конференций (мероприятия), на которых они сделаны. Выбор осуществляется в окне «Показывать». В строки запроса вводятся слова без знаков препинания, возможно использование правого усечения (символ *); по умолчанию слова связываются оператором «И», который можно поменять на «ИЛИ». Строки по умолчанию связаны оператором «И», который можно поменять на «ИЛИ» и «И НЕ».

Номер гранта И

И

Организация И

И

Название доклада И

Год: от по

Показывать по

Язык

Сортировка в

Найдено записей: 7

[CEUR Proceedings of the 21st Conference on Scientific Services & Internet \(SSI-2019\) Novorossiysk-Abrau, Russia, September 23-28, 2019.](#)

[CEUR Proceedings of the 22nd Conference on Scientific Services & Internet \(SSI-2020\) Novorossiysk-Abrau \(online\), Russia, September 21-25.](#)

[Информационные ресурсы России](#)

[Научный сервис в сети Интернет: труды XXI Всероссийской научной конференции \(23-28 сентября 2019 г., г. Новороссийск\). — М.: ИПМ им. М.В.Келдыша](#)

[Научный сервис в сети Интернет: труды XXII Всероссийской научной конференции \(21-25 сентября 2020 г., онлайн\). — М.: ИПМ им. М.В.Келдыша](#)

[Труды НИИСИ РАН. Математическое и компьютерное моделирование сложных систем: теоретические и прикладные аспекты](#)

[Цифровизация культуры и культура цифровизации: современные проблемы информационных технологий : материалы Всерос. науч. конф. \(08 октября 2020 г.\) — М. : Институт Наследия. — ISBN 978-5-86443-337-9](#)

Рис. 9. Запрос на поиск источников

Если в запросе (Рис. 9) в раскрывающемся списке «показывать» вместо «журналы / сборники» выбрать «публикации», система покажет список статей сотрудников ОИРС МСЦ, в которых даны ссылки на гранты РФФИ. Таких статей оказалось 11, они выдаются в виде списка (см. фрагмент, представленный на рис. 10).

Список публикаций, найденных в результате выполнения поискового запроса, содержит их библиографические описания и дополнительную информацию, которая включает DOI публикации, номер государственного задания, в рамках которого выполнена работа, и сведения о грантах, выделенных на исследование, отраженные в статье.

- [N. Kalenov, I. Sobolevskaya, A. Sotnikov Mathematical modeling of the processes of interdisciplinary collections formation in the digital libraries environment // CEUR Proceedings of the 21st Conference on Scientific Services & Internet \(SSI-2019\), Novorossiysk-Abrau, Russia, September 23-28, 2019, 2020. - Vol. 2543. - P. 391-398. Scopus 2-s2.0-85078449040 Гос. задание 0065-2019-0014 РФФИ 18-07-00893, 18-00-00372](#)
- [Nikolay Kalenov On the Formation of the Space of Scientific Knowledge Subject Ontology // CEUR Proceedings of the 22nd Conference on Scientific Services & Internet \(SSI-2020\) Novorossiysk-Abrau \(online\), Russia, September 21-25, 2020. - Vol. 2784. - P. 329-337. DOI: 10.51218/1613-0073-2784-329-337 Гос. задание 0580-2021-0016 РФФИ 20-07-00103](#)
- [Olga Ataeva, Nikolay Kalenov, Vladimir Serebryakov Ontological Approach to the Description of a Common Digital Space of Scientific Knowledge // CEUR Proceedings of the 22nd Conference on Scientific Services & Internet \(SSI-2020\) Novorossiysk-Abrau \(online\), Russia, September 21-25, 2020. - Vol. 2784. - P. 295-303. DOI: 10.51218/1613-0073-2784-295-303 РФФИ 20-07-00324, 18-00-00297, 18-00-00372](#)
- [Атаева О.М., Каленов Н.Е., Серебряков В.А. Об основных понятиях Единого цифрового пространства научных знаний // Научный сервис в сети Интернет: труды XXII Всероссийской научной конференции \(21-25 сентября 2020 г., онлайн\). — М.: ИПМ им. М.В.Келдыша, 2020. - С. 41-49. DOI: 10.20948/abrau-2020-18 РФФИ 20-07-00324, 18-00-00297, 18-00-00372](#)

Рис. 10. Фрагмент списка найденных публикаций

В описаниях публикаций авторы, названия публикации и названия журнала (сборника) являются активными ссылками и ведут, соответственно, на информацию об авторе (рис. 11), полный текст публикации (при соблюдении авторских прав) и источник (журнал, сборник). После перехода на источник пользователь может получить список всех статей, зарегистрированных в системе и опубликованных в данном источнике. Следует отметить, что для получения полного текста публикации (переход по ссылке на названии публикации) пользователю необходимо ввести логин и пароль, запрашиваемые системой.

Каленов Н.Е.

ORCID: [0000-0001-5269-0988](#) РИНЦ: [1494](#) Scopus: [56439969500](#) WoS: [F-4801-2013](#)

[Научно-исследовательский институт системных исследований \(НИИСИ РАН\)](#), [Межведомственный суперкомпьютерный центр Российской академии наук \(МСЦ РАН\)](#), [Отдел информационных ресурсов и систем \(ОИРС\)](#),

[Библиотека по естественным наукам РАН \(БЕН РАН\)](#)

Эквивалентные записи:

Kalenov N.
N. Kalenov
N.E. Kalenov
Nikolay Kalenov
Николай Каленов

Публикации: 493

[1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) ...

- [Николай Каленов Революция в оцифровке научных журналов или?.. // Наука. Троицкий вариант, 2021. - № 1 \(320\). - С. 15.](#)
- [Каленов Н.Е., Соболевская И.Н., Сотников А.Н. Единое цифровое пространство научных знаний в мировом информационном пространстве // Информационное общество, 2021. - № 1. - С. 30-41. DOI: 10.51218/1605-9921-2021-1-30-41 Гос. задание 065-2019-0014](#)
- [Каленов Н.Е., Савин Г.И., Сотников А.Н. Архитектура единого цифрового пространства научных знаний // Единое цифровое пространство научных знаний: проблемы и решения : сборник научных трудов / под ред. Н. Е. Каленова, А. Н. Сотникова — Москва : Берлин : Директмедиа Паблишинг, 2021. - С. 7-16. DOI: 10.51218/978-5-4499-1905-2-2021-7-16 Гос. задание 0580-2021-0016 РФФИ 20-07-00773](#)
- [Атаева О.М., Каленов Н.Е., Серебряков В.А. Онтологический подход к описанию единого цифрового пространства научных знаний // Электронные библиотеки, 2021. - Т. 24. - № 1. - С. 3-19. DOI: 10.26907/1562-5419-2021-24-1-3-19 РФФИ 20-07-00324, 18-00-00297, 18-00-00372](#)

Рис. 11. Пример показа информации о персоне

При переходе по ссылке на фамилию автора система показывает пользователю все эквивалентные записи для данной персоны, а также ее ORCID и идентификаторы в системах, РИНЦ, Scopus, WoS. Выдаются названия всех организаций,

связанных с персоной, которые были зарегистрированы в системе вместе с ее публикациями и докладами. Далее система указывает общее количество публикаций данной персоны, зарегистрированных в системе, и показывает их описания (см. рис. 11). В том случае, если у персоны указаны несколько организаций, их названия являются активными ссылками. Переход по ссылке выбранной организации приведет к показу публикаций данной персоны, в которых в качестве аффилиации автора указана данная организация.

В описаниях публикаций помимо авторов, названий статей и названий источников также активными ссылками являются DOI и номера государственных заданий. Ссылка на идентификатор DOI позволяет открыть страницу с описанием статьи (или с ее полным текстом).

После перехода по ссылке с номером государственного задания система покажет все публикации, исследования в которых выполнены в рамках данного государственного задания. На рис. 12 показан список публикаций, который формирует система в результате перехода по ссылке с номером государственного задания 065-2019-0014.

Государственное задание № 065-2019-0014

Публикации: 7

- Каленов Н.Е., Соболевская И.Н., Сотников А.Н. Единое цифровое пространство научных знаний в мировом информационном пространстве // Информационное общество, 2021. - № 1. - С. 30-41. DOI: 10.51218/1605-9921-2021-1-30-41 Гос. задание 065-2019-0014
- Igor Fiodorov, Alexander Sotnikov, Yury Telnov, Nixon Muganda Ochara Improving Business Processes Efficiency and Quality by Using BPMS // CEUR Workshop Proceedings of the 10th International Scientific and Practical Conference named after A. I. Kitov "Information Technologies and Mathematical Methods in Economics and Management (IT&MM-2020)", October 15-16, 2020, Moscow, Russia, 2021. - Vol. 2830. - P. 327-336. DOI: 10.51218/1613-0073-2830-327-336 Scopus: 2-s2.0-85103226250 Гос. задание 065-2019-0014 РФФИ 19-07-01137 А, 19-57- 60004/19
- Власова С.А., Каленов Н.Е., Сотников А.Н. Web-ориентированная система формирования контента единого цифрового пространства научных знаний // Программные продукты и системы, 2020. - № 3. - С. 365-374. DOI: 10.15827/0236-235X.131.365-374 Гос. задание 065-2019-0014
- Власова С.А. Автоматизированная система поддержки базы данных научных трудов сотрудников академических учреждений // Информационные ресурсы России, 2020. - № 5. - С. 29-31. DOI: 10.51218/0204-3653-2020-5-29-31 Гос. задание 065-2019-0014
- Власова С.А., Каленов Н.Е. Информационная система "Научные труды сотрудников академических учреждений" // Научный сервис в сети Интернет: труды XXII Всероссийской научной конференции (21-25 сентября 2020 г., онлайн). — М.: ИПМ им. М.В.Келдыша, 2020. - С. 162-165. DOI: 10.20948/abrau-2020-8 Гос. задание 065-2019-0014
- Власова С.А., Каленов Н.Е., Сотников А.Н. Экспертный подход к формированию контента электронных библиотек // Информационные ресурсы России, 2020. - № 3. - С. 2-7. DOI: 10.51218/0204-3653-2020-3-2-7 Гос. задание 065-2019-0014
- Власова С.А., Каленов Н.Е., Костюк К.Н. Автоматизированная система проведения сетевых конкурсов на лучшую электронную публикацию // Научные и технические библиотеки, 2020. - № 10. - С. 119-138. DOI: 10.33186/1027-3689-2020-10-119-138 WOS: 000596862700007 Гос. задание 065-2019-0014

Библиографическое описание
Формат CSV
Все данные
Все данные

Выгрузка записей Выделить все Отменить выбор

Рис. 12. Список публикаций с номером государственного задания 065-2019-0014

Система обеспечивает возможность выгрузки описаний публикаций, необходимых пользователю. Для этого пользователь отмечает «галочками» нужные

ему записи и выбирает формат выгрузки: библиографическое описание, формат CSV, все данные (см. рис. 12).

При выборе формата выгрузки «библиографическое описание» записи публикаций будут выданы в виде библиографических описаний стандартного вида и могут быть использованы для внесения в список пристатейной библиографии путем простого копирования. В случае выбора формата CSV данные публикаций будут загружены в EXEL-таблицу, которую пользователь может скачать на свой ПК. Наконец, формат «все данные» предоставляет возможность выгрузки полной информации о выбранных публикациях. Пример выгрузки полной информации о публикации приведен на рис. 13. Помимо элементов библиографического описания пользователю предоставляются следующие данные: вид публикации, язык публикации, ISSN (ISBN) источника, DOI, номера грантов, государственного задания, адрес полного текста. Кроме того, у каждого автора указывается его аффилиация для данной публикации.

Власова С.А.

Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)

Каленов Н.Е.

Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)

Костюк К.Н.

Издательство "Директ-Медиа" (Директ-Медиа)

Название публикации: Автоматизированная система проведения сетевых конкурсов на лучшую электронную публикацию

Вид публикации: статья в журнале

Источник: Научные и технические библиотеки

ISSN: 1027-3689 (Print), 2686-8601 (Online)

Год: 2020

Номер: 10

Страницы: 119-138

Язык публикации: русский

Идентификаторы во внешних базах данных: DOI: 10.33186/1027-3689-2020-10-119-138 WOS: 000596862700007

Номер государственного задания: 065-2019-0014

Адрес полного текста: <http://dirsmc.ru/bd/publ/HH2188.pdf>

Рис. 13. Пример полной информации о публикации

Описания докладов, найденных в результате выполнения поискового запроса, содержат: авторов доклада, название доклада, описание мероприятия, на

котором сделан доклад (рис. 14).

- Каленов Н.Е.** Об одном подходе к формированию предметных онтологий различных областей науки. [XXII Всероссийская научная конференция "Научный сервис в сети Интернет". 21.09 - 25.09.2020 г., Россия, онлайн](#)
[Презентация доклада](#) [Видеозапись выступления](#)

- Власова С.А.**, **Каленов Н.Е.** Информационная система «Научные труды сотрудников академических учреждений». [XXII Всероссийская научная конференция "Научный сервис в сети Интернет". 21.09 - 25.09.2020 г., Россия, онлайн](#)
[Презентация доклада](#) [Видеозапись выступления](#)

- Власова С.А.** Типовая система поддержки базы данных трудов научных сотрудников. [Всероссийская научная конференция "Единое цифровое пространство научных знаний: проблемы и решения". 10.11 - 13.11.2020 г., Россия, г. Москва](#)
[Презентация доклада](#) [Видеозапись выступления](#)

Рис. 14. Пример описаний докладов

В описании доклада фамилия докладчика выделяется жирным шрифтом. Ссылка от названия мероприятия обеспечивает переход на сайт мероприятия. Фамилии авторов представляют собой активные ссылки, переход по которым обеспечит выдачу всех докладов данного автора, зарегистрированных в системе. Если в системе введены адреса презентации и видеозаписи выступления, то под описанием доклада будут расположены соответствующие ссылки. Переход по ссылке «Презентация доклада» обеспечит выгрузку презентации на ПК пользователя. Ссылка «Видеозапись выступления» запускает видеофайл в браузере пользователя.

Так же, как и в случае с публикациями, система позволяет выгружать выбранные записи докладов в различных форматах. На рис. 15 приведен пример выгрузки полной информации о докладе. Выдаются все данные доклада, зарегистрированные в системе, а также все данные мероприятия, на котором сделан доклад.

Каленов Н.Е. - докладчик

Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)

Сотников А.Н.

Научно-исследовательский институт системных исследований (НИИСИ РАН). Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН). Отдел информационных ресурсов и систем (ОИРС)

Название доклада: Единое цифровое пространство знаний как интегратор библиотечной архивной и музейной информации

Вид доклада: приглашенный

Язык: русский

Адрес презентации: <http://dirsmc.ru/bd//bd/pr/2021/64.pptx>

Мероприятие: Всероссийский библиотечный конгресс: XXV Юбилейная Ежегодная Конференция Российской библиотечной ассоциации. 17.05 - 21.05.2021 г., Россия, г. Петрозаводск

Вид мероприятия: конгресс

Статус мероприятия: российское

Сайт мероприятия: <http://www.rba.ru/activities/conference/conf-2020/index/>

Рис. 15. Пример полной информации о докладе

ЗАКЛЮЧЕНИЕ

Представленная автоматизированная система регистрации результатов интеллектуальной деятельности сотрудников научного учреждения в настоящее время функционирует в технологическом режиме в МСЦ РАН. На начало октября 2021 года в системе зарегистрированы 338 персон из 51 организации; 802 статьи, опубликованные в 362 изданиях; 58 докладов, сделанных на 34 мероприятиях в 2018–2021 гг.

Система может быть использована в любой научной организации и предоставлять неоценимую помощь в работе как научных сотрудников, так и администрации организации. Научный сотрудник в поисковом модуле системы сможет найти свои публикации и доклады за заданный период времени и использовать эти данные для отчетов или составления пристатейной библиографии. Ученый секретарь организации может получить исчерпывающую информацию о публикациях и выступлениях сотрудников за заданный промежуток времени или данные о статьях, выполненных в рамках конкретного государственного задания или гранта. Система предоставляет возможность получения всей необходимой информации для формирования годового отчета организации в Министерство науки и высшего образования.

Работа выполнена в МСЦ РАН – филиале ФГУ ФНЦ НИИСИ РАН в рамках государственного задания № 0580-2021-0014.

СПИСОК ЛИТЕРАТУРЫ

1. *Бескаравайная Е.В., Довбня Е.В., Захарова С.С.* Проблемно-ориентированные коллекции. Формирование и анализ на примере базы данных трудов сотрудников Института биофизики клетки // Библиография. 2008. № 4. С. 30–36.
2. *Захарова С.С., Гуреева Ю.А.* Научные публикации: от картотеки трудов до библиографических профилей // Библиосфера. 2017. № 2. С. 85–89. <https://doi.org/10.20913/1815-3186-2017-2-85-89>.
3. *Левченко О.И., Соловьев А.В.* Формирование базы данных публикаций сотрудников Института физики твердого тела РАН // Информационное обеспечение науки: новые технологии: Сборник научных трудов. М.: БЕН РАН, 2015. С. 215–221.
4. *Рогозникова О. А., Данилин М. В.* Интеграция базы данных публикаций организации с индексами научного цитирования: реализация средствами САБ ИРБИС-64 // Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса: материалы Международной конференции. 2015.
5. *Мазов Н.А., Гуреев В.Н.* Библиографическая база данных трудов сотрудников организации: цели, функции, сфера использования в наукометрии // Вестник Дальневосточной государственной научной библиотеки. 2016. Вып. 2 (71). С. 84–87.
6. *Ковязина Е.В.* БД трудов сотрудников как средство учета и продвижения научных публикаций // Труды ГПНТБ СО РАН. 2017. № 12. С. 336–343.
7. *Панкратов И.А., Ратушный А.В.* Проектирование информационной системы для хранения информации о научных публикациях // Вестник молодежной науки России. 2019. № 4. С. 31.
8. ORCID – Open Researcher and Contributor ID. URL: <https://orcid.org> (дата обращения: 10.10.2021).
9. Research Organization Registry Community. URL: <https://ror.org> (дата обращения: 10.10.2021).

10. База данных публикаций ИФТТ.
URL: http://www.issp.ac.ru/libcatm/publications_m.php (дата обращения: 10.10.2021).
 11. Публикации сотрудников УлГУ.
URL: https://www.ulsu.ru/ru/page/page_1777/ (дата обращения: 10.10.2021).
 12. Публикации сотрудников МИАН.
URL: [https://www.mi-ras.ru/index.php?c=mianpubs&l=0&jrnfilters\[\]=jher](https://www.mi-ras.ru/index.php?c=mianpubs&l=0&jrnfilters[]=jher) (дата обращения: 10.10.2021).
 13. ИСТИНА (руководство пользователя).
URL: https://docs.istina.msu.ru/getting_started/main.html (дата обращения: 10.10.2021).
 14. *Власова С.А., Каленов Н.Е.* Новые подходы к формированию баз данных публикаций сотрудников академических учреждений // Научные труды Института рукописей Национальной академии наук Азербайджана. 2018. № 2 (7). С. 85–94.
 15. *Власова С.А.* Автоматизированная система поддержки корпоративной базы данных научных публикаций // Программные продукты, системы и алгоритмы. 2018. Вып. 2. С. 42–46. <https://doi.org/10.15827/2311-6749.27.311>
 16. *Власова С.А., Каленов Н.Е.* Информационная система «Научные труды сотрудников академических учреждений» // В сборнике: Научный сервис в сети Интернет труды XXII Всероссийской научной конференции. ИПМ им. М.В. Келдыша. 2020. С. 152–165. <https://doi.org/10.20948/abrau-2020-8>
 17. *Svetlana Vlasova, Nikolay Kalenov.* Information System for Registering the Result of Scientific Institution Employees' Intellectual Activity // CEUR Workshop Proceedings. 2020. Vol. 2784. P. 283–294. <https://doi.org/10.51218/1613-0073-2784-283-294>
-

DEVELOPMENT OF THE INFORMATION SYSTEM FOR REGISTERING THE RESULT OF SCIENTIFIC INSTITUTION' EMPLOYEES INTELLECTUAL ACTIVITY

S. A. Vlasova¹ [0000-0003-1533-5850], N. E. Kalenov² [0000-0001-5269-0988]

¹⁻² Joint Supercomputer Center of the Russian Academy of Sciences – JSCC

¹*vlas.svetlana2013@yandex.ru*, ²*nekalenov@mail.ru*

Abstract

The article describes a Web-system developed by the authors that implements services related to the formation and provision of multifaceted information about the results of scientific activities (publications, copyright certificates and reports at scientific events) of employees of an organization or a group of organizations. The system is focused both on the end user interested in obtaining specific data, and on the administrative staff, who generates reporting materials for the parent organization. The information base of the system contains metadata on the following classes of objects: persons (authors), organizations and their subdivisions; publications at analytical, monographic and summary levels; copyright certificates; scientific events (conferences, symposia, seminars); reports. The system includes two modules – an administrative one (intended for entering and editing data) and a user one, which is a special search engine that searches for information, visualizes it, provides navigation among related resources and exports data. A distinctive feature of the system is the introduced concept of “equivalent” objects. Objects are considered equivalent if they are represented in the system by different metadata, but referring to the same physical entity. Such objects are “persons” corresponding to one author with different spellings of the surname in the bibliographic descriptions of publications; organizations with different variants of names; articles published unchanged in various languages. In accordance with modern requirements for reporting on publications, the system reflects the sources of research funding, as well as the affiliations indicated in the articles for each author.

Keywords: *scientific works, scientific activity, automated system, database, management reports, network technologies.*

REFERENCES

1. *Beskaravajnaja E.V., Dovbnja E.V., Zaharova S.S.* Problemnoorientirovannye kollekcii. Formirovanie i analiz na primere bazy dannyh trudov sotrudnikov Instituta biofiziki kletki // Bibliografija. 2008. № 4. S. 30–36.
2. *Zaharova S.S., Gureeva Ju.A.* Nauchnye publikacii: ot kartoteki trudov do bibliograicheskikh profilej // Bibliosfera. 2017. № 2. S. 85–89.
<https://doi.org/10.20913/1815-3186-2017-2-85-89>.
3. *Levchenko O.I., Solov'ev A.V.* Formirovanie bazy dannyh publikacij sotrudnikov Instituta fiziki tverdogo tela RAN // Informacionnoe obespechenie nauki: novye tehnologii: Sbornik nauchnyh trudov. M.: BEN RAN. 2015. S. 215–221.
4. *Rogoznikova O.A., Danilin M.V.* Integracija bazy dannyh publikacij organizacii s indeksami nauchnogo citirovanija: realizacija sredstvami SAB IRBIS-64 // Biblioteki i informacionnye resursy v sovremennom mire nauki, kul'tury, obrazovanija i biznesa: materialy Mezhdunarialy konferencii. 2015.
5. *Mazov N.A., Gureev V.N.* Bibliograficheskaja baza dannyh trudov sotrudnikov organizacii: celi, funkcii, sfera ispol'zovanija v naukometrii // Vestnik Dal'nevostochnoj gosudarstvennoj nauchnoj biblioteki, 2016. Vyp. 2 (71). S. 84–87.
6. *Kovjazina E.V.* BD trudov sotrudnikov kak sredstvo ucheta i prodvizhenija nauchnyh publikacij // Trudy GPNTB SO RAN. 2017. № 12-2. S. 336–343.
7. *Pankratov I.A., Ratushnyj A.V.* Proektirovanie informacionnoj sistemy dlja hranenija informacii o nauchnyh publikacijah // Vestnik molodezhnoj nauki Rossii. 2019. № 4. S. 31.
8. ORCID – Open Researcher and Contributor ID. URL: <https://orcid.org> (accessed 10 October 2021).
9. Research Organization Registry Community. URL: <https://ror.org> (accessed 10 October 2021).
10. Baza dannyh publikacij IFTT. URL: http://www.issp.ac.ru/libcatm/publications_m.php (accessed 10 October 2021).
11. Publikacii sotrudnikov UIGU.
URL: https://www.ulsu.ru/ru/page/page_1777/ (accessed 10 October 2021).
12. Publikacii sotrudnikov MIAN.
URL: [https://www.mi-ras.ru/index.php?c=mianpubs&l=0&jrnfilters\[\]=jhep](https://www.mi-ras.ru/index.php?c=mianpubs&l=0&jrnfilters[]=jhep) (accessed

10 October 2021).

13. ISTINA (rukovodstvo pol'zovatelja).

URL: https://docs.istina.msu.ru/getting_started/main.html (accessed 10 October 2021).

14. *Vlasova S.A., Kalenov N.E.* Novye podhody k formirovaniyu baz dannyh publikacij sotrudnikov akademicheskikh uchrezhdenij // Nauchnye trudy Instituta rukopisej Nacional'noj akademii nauk Azerbajdzhana. 2018. № 2 (7). S. 85–94.

15. *Vlasova S.A.* Avtomatizirovannaja sistema podderzhki korporativnoj bazy dannyh nauchnyh publikacij // Programmnye produkty, sistemy i algoritmy. 2018. Vyp. 2. S. 42–46. <https://doi.org/10.15827/2311-6749.27.311>

16. *Vlasova S.A., Kalenov N.E.* Informacionnaja sistema "Nauchnye trudy sotrudnikov akademicheskikh uchrezhdenij" // V sbornike: Nauchnyj servis v seti Internet trudy XXII Vserossijskoj nauchnoj konferencii. IPM im. M.V. Keldysha. 2020. S. 152–165. <https://doi.org/10.20948/abrau-2020-8>

17. *Svetlana Vlasova, Nikolay Kalenov.* Information System for Registering the Result of Scientific Institution Employees' Intellectual Activity // CEUR Workshop Proceedings. 2020. Vol. 2784. P. 283–294. <https://doi.org/10.51218/1613-0073-2784-283-294>

СВЕДЕНИЯ ОБ АВТОРАХ



ВЛАСОВА Светлана Александровна – ведущий научный сотрудник Межведомственного Суперкомпьютерного Центра РАН – филиала Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», кандидат технических наук.

Svetlana Aleksandrovna VLASOVA – Leading Researcher of Joint Super Computer Center of the Russian Academy of Sciences – Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences”, Candidate of Technical Sciences.

email: vlas.svetlana2013@yandex.ru;

ORCID: 0000-0003-1533-5850



КАЛЕНОВ Николай Евгеньевич – главный научный сотрудник Межведомственного Суперкомпьютерного Центра РАН – филиала Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», доктор технических наук, профессор.

Nikolay Evgenyevich KALENOV – Chief Researcher of Joint Super Computer Center of the Russian Academy of Sciences – Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences”, Doctor of Technical Sciences, Professor.

email: nekalenov@mail.ru;

ORCID: 0000-0001-5269-0988

Материал поступил в редакцию 19 октября 2021 года

УДК 01

ПРОСПЕКТИВНАЯ БИБЛИОГРАФИЯ

М. М. Горбунов-Посадов^[0000-0002-7044-8287]

Институт прикладной математики им. М.В. Келдыша РАН, Миусская пл., 4,
Москва, 125047

gorbunov@keldysh.ru

Аннотация

Перспективная библиография, т. е. список ссылающихся публикаций — динамически составляемый перечень работ, вышедших вслед за рассматриваемым изданием и содержащих ссылку на него. Ведущие зарубежные издатели научных журналов, как правило, размещают списки ссылающихся публикаций на веб-страницах своих статей. В нашей стране размещение списков ссылающихся публикаций пока еще не стало нормой, однако постепенно получает распространение. Рассматриваются реализации списка ссылающихся публикаций в ряде характерных проектов. Обсуждается используемая терминология.

Ключевые слова: *список ссылающихся публикаций, цитирующие публикации, обратный библиографический список.*

ВВЕДЕНИЕ

Традиционным атрибутом любого научного издания является библиография, демонстрирующая, на какие работы опиралось это издание. Библиография несет очевидную этическую нагрузку: с ее помощью авторы как бы приносят благодарность своим предшественникам, работы которых послужили фундаментом для подготовки нового издания. Библиография интересна не только историку науки, но и, конечно же, обычному читателю, позволяя ему углубиться в изучение предмета, узнать важные обстоятельства, подробные сведения о которых по тем или иным причинам не вошли непосредственно в текст нового издания, но добросовестно отражены в его библиографическом списке.

Однако существенно больший интерес для читателя представляет список ссылающихся публикаций — перечень работ, вышедших вслед за данным изданием и содержащих ссылку на него. Ссылающиеся публикации позволяют узнать,

в каком направлении развиваются исследования, представленные в исходном издании. Иначе говоря, если обычный (ретроспективный) библиографический список смотрит как бы назад, в прошлое, то проспективный список, т. е. список ссылающихся публикаций, смотрит вперед, в будущее, охватывая новейшие работы в рассматриваемой области науки.

Механизмы формирования обычного библиографического списка и списка ссылающихся публикаций различны, хотя и тесно связаны между собой. Обычный, традиционный библиографический список составляет автор, и делает он это сейчас практически так же, как делали до него сотни лет назад. И напротив, списки ссылающихся публикаций генерируют библиографические базы с использованием современных информационных технологий.

Публикация научной работы в настоящее время подразумевает размещение сведений о ней в открытом доступе в интернете, а также передачу метаатрибутов работы в библиографические базы. В данном случае существенно, что среди размещаемых и передаваемых данных присутствует библиографический список работы. Составляющие список библиографические записи попадают в базы данных, поддерживаемые Google Scholar и библиографическими базами. На основе таких баз данных не только вычисляются многочисленные библиометрические оценки, интересующие в первую очередь чиновников от науки, но и генерируются списки ссылающихся публикаций, чрезвычайно полезные для массового читателя.

Если библиография, т. е. обычный библиографический список, может с успехом существовать и в печатном издании, и в онлайн, то список ссылающихся публикаций — конструкция по существу исключительно онлайн-овая, динамическая, формируемая «на лету» на экране онлайн-ового читателя. Список ссылающихся публикаций всякий раз приносит читателю свежее текущее состояние накапливаемых библиографических данных. Здесь практически сразу отражаются появляющиеся в интернете и индексируемые в библиографических базах публикации, содержащие в своей библиографии ссылку на рассматриваемую научную работу.

Терминология в данной области пока не устоялась. Наряду с используемым здесь термином «список ссылающихся публикаций» встречаются также и «обратный библиографический список», и «список цитирующих публикаций». Однако оба этих оборота имеют определенные слабости. «Обратный» список очевидным образом удачно противопоставляется «прямому» — обычной библиографии, но в слове «обратный» в данном случае иногда слышится прямо противоположный по смыслу оттенок: «повернутый назад, в прошлое». «Цитирующая публикация» русскоязычным читателем может восприниматься слишком буквально — как публикация, включающая цитаты из указанного источника, в то время как такое явное механическое копирование частей текста встречается в научных работах относительно редко. Вместе с тем в англоязычной литературе для обозначения интересующего нас списка чаще всего используются именно термины “cited-by” и “citing”.

Более точно и рельефно противопоставляются традиционная библиография и ссылающиеся публикации при использовании соответствующих терминов «ретроспективный» и «проспективный» (библиографический список). К сожалению, в русскоязычной научной лексике слово «проспективный», т. е. направленный в будущее, относящийся к движению вперед, встречается редко, а в контексте библиографии пока еще просто не используется. Данная статья получила название «проспективная библиография» именно в надежде привлечь внимание читателя к этому удачному термину.

Заинтересованность онлайн-читателя в аппарате списков ссылающихся публикаций не вызывает сомнений. Большинство ведущих зарубежных издателей научных журналов уже размещают в том или ином виде списки ссылающихся на веб-страницах публикуемых статей. У нас в стране списки ссылающихся встречаются пока нечасто. В настоящей работе рассматриваются некоторые получившие известность списки ссылающихся публикаций в крупных зарубежных проектах, а также реализация таких списков в онлайн-версии издания «Препринты ИПМ им. М.В. Келдыша».

GOOGLE SCHOLAR

Наиболее широко известна и популярна реализация списка ссылающихся публикаций, предлагаемая Google Scholar. Роботы Google Scholar систематически

бороздят просторы интернета в поисках научных работ. Каждая найденная научная публикация подробно разбирается, в частности, в ней вычленяется и анализируется библиографический список. Записи библиографического списка сохраняются в Google Scholar, и далее построенная из них база данных используется, в частности, для построения списков ссылающихся публикаций.

Чтобы получить список ссылающихся публикаций, в поле поиска Google Scholar необходимо указать либо адрес (URL) интересующей нас публикации, либо записать в это поле авторов и название публикации. И в том, и в другом случае Google Scholar выдает результат, подобный представленному на рис. 1.

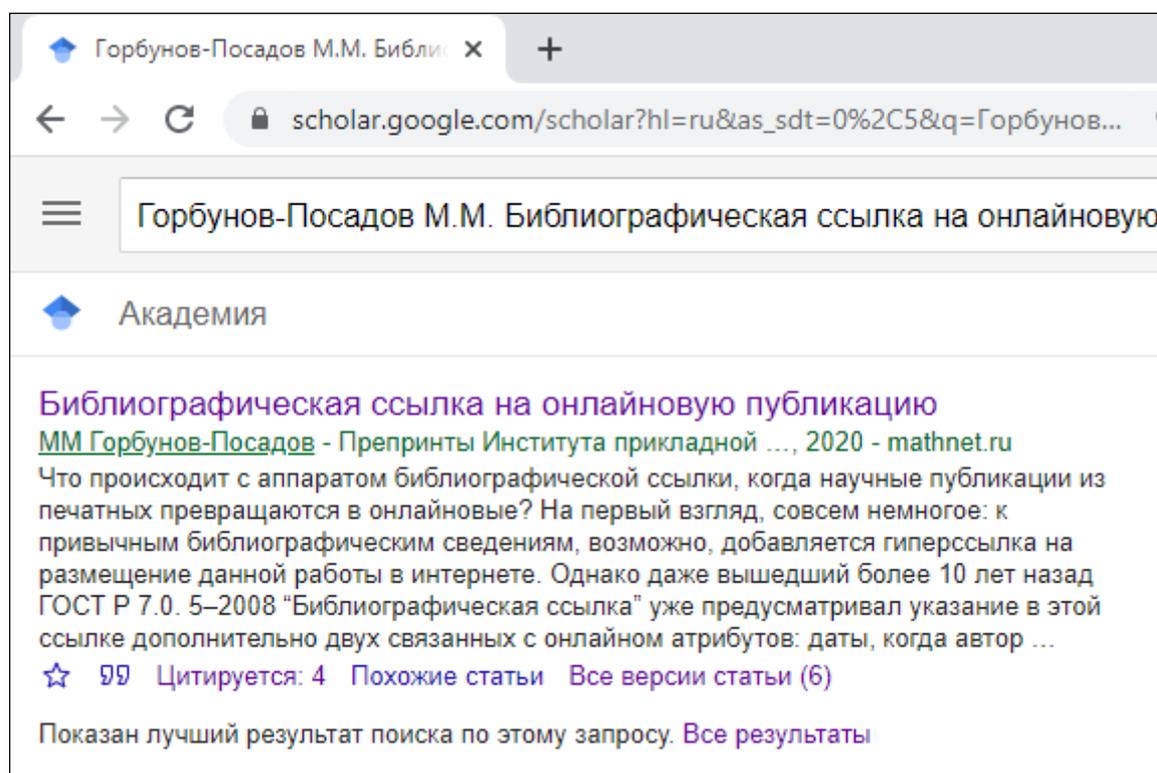


Рис. 1. Результат поиска в scholar.google.com строки «Горбунов-Посадов М.М. Библиографическая ссылка на онлайн-публикацию»

Теперь, чтобы увидеть список публикаций, ссылающихся на заданную, надо кликнуть на расположенном в предпоследней строке поле «Цитируется: 4». Первые две записи результата такого обращения представлены на рис. 2.

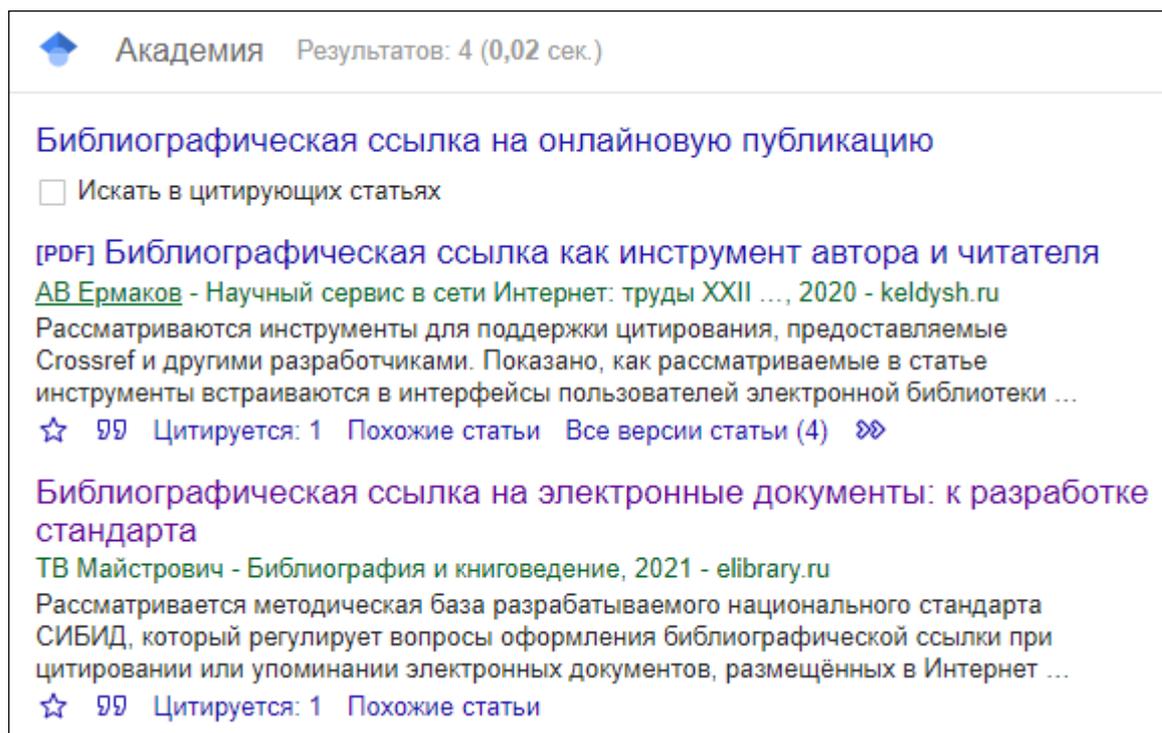


Рис. 2. Список ссылающихся публикаций, построенный Google Scholar.

Желание познакомиться со списком ссылающихся обычно возникает у читателя при работе с онлайн-публикацией. Поэтому разработчики современной среды, обслуживающей онлайн-публикации, стремятся сократить число манипуляций, необходимых в этой среде для получения списка ссылающихся публикаций. Наиболее естественное и удобное в данном случае решение — размещение заблаговременно сформированной ссылки на формируемый список непосредственно в карточке публикации. К счастью, Google Scholar позволяет статически сформировать такую ссылку. Среди отечественных онлайн-хранилищ научных публикаций этой полезной возможностью воспользовались, в частности, портал MathNet.ru, библиотека eLibrary.ru и Препринты ИПМ им. М.В. Келдыша (рис. 3).

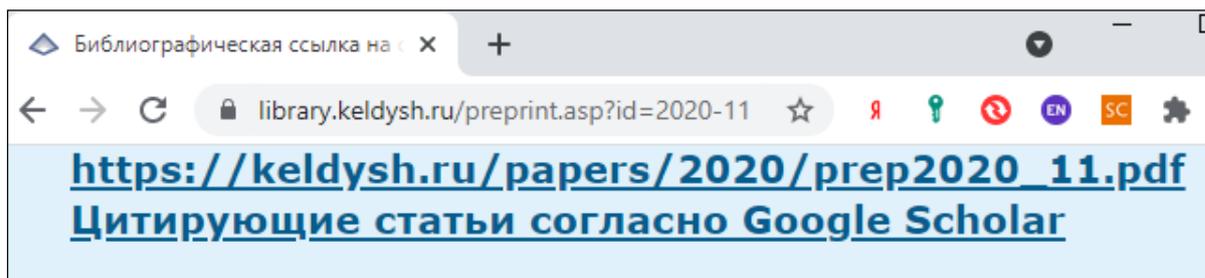


Рис. 3. Ссылка, ведущая на формирование списка ссылающихся публикаций, на сайте Препринтов ИПМ им. М.В. Келдыша.

За рубежом создатели библиографических систем нередко идут еще дальше. Если онлайн-публикация представлена в формате HTML, то путь к просмотру списка ссылающихся можно дополнительно сократить, встроив этот список непосредственно в страницу полного текста публикации. В самом деле, для читателя список ссылающихся ничуть не менее интересен, чем традиционный библиографический список, и поэтому его доступность естественно следует обеспечить не хуже. Для этого потребуется разместить список ссылающихся либо до, либо после традиционного списка, или же, при достаточной ширине экрана, вынести его в параллельную колонку.

ELIBRARY.RU

Помимо упомянутой выше размещаемой на странице публикации отсылки к списку ссылающихся, формируемому Google Scholar, научная электронная библиотека eLibrary формирует и свой список, опирающийся на данные, хранящиеся в этой библиотеке. Российскому ученому список ссылающихся из eLibrary особенно интересен, поскольку в фондах eLibrary хранится богатейшая коллекция сведений именно о российских научных публикациях. Увидеть список ссылающихся из eLibrary (рис. 4), вообще говоря, несложно: сначала необходимо с помощью поиска добраться до сформированной eLibrary страницы интересующей публикации, а затем перейти по расположившейся справа на этой странице ссылке «Список статей в РИНЦ, цитирующих данную».

СПИСОК СТАТЕЙ, ЦИТИРУЮЩИХ ДАННУЮ ПУБЛИКАЦИЮ



ИНТЕРНЕТ-АКТИВНОСТЬ КАК ОБЯЗАННОСТЬ УЧЕНОГО
Горбунов-Посадов М.М.
 Информационные технологии и вычислительные системы. 2007. № 3. С. 88-93.

Журнал

Автор Год Название

Сортировка по дате выпуска Порядок по возрастанию Очистить Выборка

	Публикация	Цит.
1	<p>ИНФОРМАЦИОННОЕ ПРОСТРАНСТВО РОССИЙСКОЙ АКАДЕМИЧЕСКОЙ НАУКИ КАК ИНТЕГРИРОВАННАЯ ПРОФЕССИОНАЛЬНО-ИНТЕЛЛЕКТУАЛЬНАЯ СРЕДА: ПЕРСПЕКТИВЫ ФОРМИРОВАНИЯ И УПРАВЛЕНИЯ <i>Соколова М.Е.</i> В сборнике: Концепция "общества знания" в современной социальной теории. Сб. науч. тр. Сер. "Теория и история социологии" Центр социал. науч.-информ. исслед. Отд. социологии и социал. психологии; Отв. ред. Д.В. Ефременко. МОСКВА, 2010. С. 134-158.</p>	9
2	<p>МЕТОДОЛОГИЧЕСКИЕ ВОПРОСЫ СОВЕРШЕНСТВОВАНИЯ НАУЧНОЙ КОММУНИКАЦИИ С ИСПОЛЬЗОВАНИЕМ СОВРЕМЕННЫХ ИНФОРМАЦИОННО-КОМУНИКАЦИОННЫХ ТЕХНОЛОГИЙ <i>Балл Г.А., Мединцев В.А.</i> Информационные технологии и средства обучения. 2011. Т. 22. № 2. С. 2.</p>	1
3	<p>КИБЕРСОЦИАЛИЗАЦИЯ ЧЕЛОВЕКА. ОТ НОМО SAPIENS'A ДО НОМО CYBERUS'A <i>Плешаков В.А.</i> Москва, 2012.</p>	101

Рис. 4. Список ссылающихся публикаций в eLibrary.ru.

В то же время, как уже упоминалось, желание ознакомиться со списком ссылающихся обычно возникает у онлайн-читателя в момент прочтения онлайн-публикации. Если в этот момент список ссылающихся оказывается «под рукой», т. е. ссылка на список размещена непосредственно на веб-странице публикации, то читатель вполне может этой ссылкой заинтересоваться и перейти по ней. Если же для доступа к списку требуется выполнить весьма трудоемкие манипуляции — авторизоваться в библиографической системе, разыскать в ней заинтересовавшую публикацию и т. д. — в таком случае лишь редкий читатель будет готов потратить здесь свое время и силы.

Тут eLibrary безнадежно проигрывает Google Scholar. Если непосредственную ссылку на список ссылающихся, генерируемый Google Scholar, разместить на веб-странице публикации не составляет труда, то построить подобную ссылку для списка ссылающихся, формируемого eLibrary, не удастся. Дело в том, что список ссылающихся в eLibrary не имеет постоянного (статического) адреса (URL).

Надежды онлайн-библиотек научных публикаций одно время связывались с недавно построенным API для доступа к отдельным данным, хранящимся в eLibrary, однако в состав этого API доступ к списку ссылающихся публикаций, к сожалению, не вошел.

SCOPUS

В отличие от eLibrary наиболее крупные международные библиографические базы все же, как правило, так или иначе предоставляют возможность разместить непосредственно на веб-странице публикации список ссылающихся или же гиперссылку на него. На рис. 5 представлен список ссылающихся публикаций для статьи [1] из журнала "Information Processing & Management", составленный на основе данных библиографической базы Scopus и размещенный на веб-странице этой статьи.



Рис. 5. Scopus: список ссылающихся публикаций для [1].

На рис. 5 несколько неожиданно появление третьей позиции: там упоминается статья, взятая не из основного хранилища Scopus, а с сервера препринтов arXiv.org. Дело в том, что в 2021 году Scopus при составлении списка ссылающихся публикаций начал [2] включать в рассмотрение наряду со своими основными данными еще и статьи, размещенные на популярных серверах препринтов arXiv, bioRxiv, ChemRxiv и medRxiv. Scopus справедливо отмечает, что для читателя особый интерес представляют именно самые свежие публикации, сейчас доступные прежде всего на этих серверах.

Повышенный интерес к серверам препринтов появился в связи с массовым размещением там публикаций, посвященных пандемии Ковид-19. Для таких публикаций ключевую роль играла скорость появления их в общедоступном онлайн, а тут препринтам, как известно, просто нет равных. В период пандемии в форме препринта выходило около половины всех публикаций, посвященных коронавирусу.

CROSSREF

Списки ссылающихся, составляемые Crossref на основе данных о публикациях, получающих DOI, — одни из наиболее популярных и авторитетных. В частности, проспективные библиографические списки Crossref размещает на веб-страницах своих изданий издательство Springer (рис. 6).

Citation Details

Article

Norms modeling constructs of business process compliance management frameworks: a conceptual evaluation

Artificial Intelligence and Law, 2018, Volume 26, Number 3, Page 251
Mustafa Hashmi, Guido Governatori

[Read Online](#)

5 ITEMS CITE THIS ARTICLE

Page: 1

Chapter

0

CITATIONS

Specifics of Decision Making in Modern Business Systems

Svetlana V. Lobova, Anna V. Bodiako, Liudmila V. Dontsova, Yevgeniy An and Viktor N. Salin

Year: 2019, Page 21
DOI: 10.1108/978-1-78756-691-020191004

[Read Online](#)

Article

29

CITATIONS

Are we done with business process compliance: state of the art and challenges ahead

Mustafa Hashmi, Guido Governatori, Ho-Pun Lam and Moe Thandar Wynn

Journal: Knowledge and Information Systems, 2018, Volume 57, Number 1, Page 79
DOI: 10.1007/s10115-017-1142-1

[Read Online](#)

Рис. 6. Список ссылающихся публикаций Crossref на сайте Springer.

На сайте Crossref список ссылающихся публикаций можно получить только посредством обращения к API, т. е. здесь требуется написать программный код. Такой код реализован, в частности, на сайте Препринтов ИПМ им. М.В. Келдыша. Результат представлен на рис. 7.

Цитирование публикации  English

М.М. Горбунов-Посадов
Библиографическая ссылка на онлайн-публикацию
Препринты ИПМ им. М.В.Келдыша. 2020. № 11. Страниц: 26. Язык: русский
<https://doi.org/10.20948/prepr-2020-11>

Цитирований в Crossref: 2 

Conference paper
М.М. Gorbunov-Possadov
Bibliographic reference on the Internet
Издание: Proceedings of the 3rd International Conference "Futurity designing. Digital reality problems", 2020, Page 234
<https://doi.org/10.20948/future-2020-21>
Цитирований: 0

Conference paper
A.V. Ermakov
Bibliographic reference as a tool for the author and reader
Издание: Proceedings of 22nd Scientific Conference "Scientific Services & Internet – 2020", 2020, Page 268
<https://doi.org/10.20948/abrau-2020-55>
Цитирований: 1

Рис. 7. Список ссылающихся публикаций Crossref на сайте Препринтов ИПМ им. М.В. Келдыша. Калька с английского: «цитирующие», а не «ссылающиеся» публикации.

Однако рядовой читатель, имеющий DOI заинтересовавшей его публикации, получить соответствующий список ссылающихся все же может и без программирования. Для этого ему придется обратиться к сайтам сторонних разработчиков, например, к [3]. Там для получения списка ссылающихся публикаций достаточно ввести в соответствующее поисковое поле имеющееся значение DOI.

ДРУГИЕ ДИНАМИЧЕСКИЕ БИБЛИОГРАФИЧЕСКИЕ СПИСКИ

Наряду с рассмотренным выше списком ссылающихся публикаций известны и другие динамически составляемые библиографические списки, используемые в библиографических системах. Наиболее популярные из них — списки «созвучных», «родственных» публикаций, где собираются публикации, близкие по тематике к рассматриваемой работе.

На первых ролях сегодня здесь список «родственных», составляемый Google Scholar. Так, например, в библиографической системе mathnet.ru этот список можно получить, перейдя по ссылке “Related articles on Google Scholar”. Число таких работ, предлагаемых Google Scholar, невероятно велико, однако качество их отбора пока оставляет желать лучшего. В частности, Google Scholar в качестве одного из поисковых атрибутов использует просто фамилии авторов рассматриваемой публикации, поэтому заметную часть результатов составляют работы, не имеющие отношения к рассматриваемой теме, но написанные однофамильцами авторов.

Собирает список родственных публикаций и eLibrary. Этот список сопровождает каждую публикацию, к нему ведет ссылка под названием «Найти близкие по тематике публикации». И здесь читателю предлагают десятки тысяч ответов, качество этих ответов также не всегда радует, хотя и превосходит качество результатов Google Scholar.

За рубежом списки родственных публикаций также широко распространены и используют различные механизмы. Например, интересные результаты получаются в проекте Connected papers [4], где при поиске родственных публикаций под близостью публикаций понимается не наличие прямых библиографических связей между статьями, а близость их библиографических списков.

И все же для читателя как правило наибольший интерес представляет основной предмет данной статьи — список ссылающихся публикаций: механизм его построения понятен читателю и в то же время результаты приносят несомненную пользу. Отметим, что динамические библиографические списки с очевидными алгоритмами построения сейчас наиболее востребованы. Среди них, наряду со списком ссылающихся, необходимо отметить аппарат ORCID [5], позволяющий читателю в один клик увидеть список публикаций заинтересовавшего его автора.

ЗАКЛЮЧЕНИЕ

Список ссылающихся публикаций, вообще говоря, подстерегают опасности, подобные сыгравшим в свое время роковую роль в судьбе интересного проекта COinS [6]. Нечистоплотный автор может заметно подпортить отношение к этой конструкции, включив в библиографию своей работы ссылку на выдающуюся научную публикацию не по существу, а лишь в расчете на то, что его работу заметят за счет появления ее в широко востребованном списке ссылающихся на эту выдающуюся публикацию. К счастью, такое развитие событий сейчас представляется маловероятным.

Список ссылающихся постепенно становится обязательным элементом современной среды онлайн-публикаций. Остается только порадоваться за онлайн-читателя, обретшего столь полезный инструмент для работы с научной литературой.

СПИСОК ЛИТЕРАТУРЫ

1. *Yong Gao, Qiang Wu, Linna Zhu*. Merging the citations received by arXiv-deposited e-prints and their corresponding published journal articles: Problems and perspectives // *Information Processing & Management*. 2020. V. 57, No. 5. <https://doi.org/10.1016/j.ipm.2020.102267>
2. *R. McCullough*. Preprints are now in Scopus! // *blog.scopus*. URL: <https://blog.scopus.com/posts/preprints-are-now-in-scopus>
3. Open Ukrainian Citation Index (OUCI). URL: <https://ouci.dntb.gov.ua>
4. Connected papers. URL: <https://www.connectedpapers.com>
5. ORCID connecting research and researchers. URL: <https://orcid.org>
6. COinS (ContextObjects in Spans). URL: <https://en.wikipedia.org/wiki/COinS>

PROSPECTIVE BIBLIOGRAPHY

M. M. Gorbunov-Posadov ^[0000-0002-7044-8287]

Keldysh Institute of Applied Mathematics, Miusskaya sq., 4, Moscow, 125047, Russia

gorbunov@keldysh.ru

Abstract

The prospective bibliography, i.e., the list of referring publications— is a dynamically drawn-up list of works published after the publication under consideration and contained a reference to it. Leading foreign publishers of scientific journals, as a rule, post lists of referring publications on the web pages of their articles. In our country, the posting lists of referring publications has not yet become the norm, but it is gradually becoming widespread. The implementations of the list of referring publications in a number of characteristic projects are considered. The terminology used is discussed.

Keywords: *list of referring publications, citing publications, reverse bibliography*

REFERENCES

1. *Yong Gao, Qiang Wu, Linna Zhu*. Merging the citations received by arXiv-deposited e-prints and their corresponding published journal articles: Problems and perspectives // *Information Processing & Management*. 2020. V. 57, No. 5. <https://doi.org/10.1016/j.ipm.2020.102267>
2. *R. McCullough*. Preprints are now in Scopus! // *blog.scopus*. URL: <https://blog.scopus.com/posts/preprints-are-now-in-scopus>
3. Open Ukrainian Citation Index (OUCI). URL: <https://ouci.dntb.gov.ua>
4. Connected papers. URL: <https://www.connectedpapers.com>
5. ORCID connecting research and researchers. URL: <https://orcid.org>
6. COinS (ContextObjects in Spans). URL: <https://en.wikipedia.org/wiki/COinS>

СВЕДЕНИЯ ОБ АВТОРЕ



ГОРБУНОВ-ПОСАДОВ Михаил Михайлович – главный научный сотрудник ИПМ им. М.В. Келдыша РАН, д. ф.-м. н.

Mikhail Mikhailovich GORBUNOV-POSADOV – Keldysh Institute of Applied Mathematics, chief researcher. Current scientific interests: programs extensibility, scientific publications.

email: gorbunov@keldysh.ru;

ORCID: 0000-0002-7044-8287

Материал поступил в редакцию 16 октября 2021 года

УДК 001.812, 004.051, 004.738.5, 336.764.061.1

ПРИНЦИПЫ МУЛЬТИТРЕЙДИНГА

Ф. О. Каспаринский^[0000-0002-1048-9212]

ООО «МАСТЕР-МУЛЬТИМЕДИА», Шоссе Энтузиастов 98-3-274, Москва, 111531
felix@kasparinsky.pro

Аннотация

Современные программные и аппаратные средства предоставляют беспрецедентную свободу для разнообразной деятельности на форекс-рынках, от трейдинга до анализа дееспособности моделей нелинейных процессов в самоорганизующихся системах. Для снижения рисков и увеличения эффективности взаимодействия с инструментами фондовых рынков предлагается обеспечить вариативную адаптивность трейдинга посредством комбинирования торговых стратегий с использованием множества торговых счетов различных торговых площадок, разнообразных финансовых инструментов и Комплексных Индикаторов Тенденций изменения цен. В результате трёхлетней экспериментальной работы сформулированы и апробированы основные принципы мультитрейдинга и скомпонована информационная среда, способствующая разработке индивидуализированной торговой системы. Базовая концепция организации инфосреды мультитрейдинга: применение специализированных аппаратно-программных комплексов для стратегического анализа и прогноза изменения цен отдельного финансового инструмента, тактического выбора перспективного финансового инструмента из доступного набора и операционной деятельности с ордерами торговых счетов. Можно ожидать, что эволюция принципов мультитрейдинга приведёт к созданию аналитических систем прогнозирования кинетики неравновесного изменения характеристических параметров самоорганизующихся кооперативных систем для широкого применения в биологии, кибернетике, экономике и социальной сфере.

Ключевые слова: мультитрейдинг, трейдинг, форекс, технический анализ, инвестиции, организация работы, эффективность, финансовый рынок, осцилляции, прогноз.

ВВЕДЕНИЕ

Самоорганизация системных сервисов глобального информационного общества обеспечила свободное распространение профессионального инструментария для работы не только с информационными, но и финансовыми ресурсами. Доступная информационная среда для работы с биржевыми финансовыми инструментами создает условия для вовлечения денежных средств каждого пользователя интернета в глобальный финансовый рынок. Профессионально неподготовленные участники рыночных процессов перераспределения ресурсов дезориентируются в многообразии финансовых инструментов, биржевых посредников, торговых платформ и способах их использования. Бессистемное комбинирование и некомпетентное использование компонентов профессиональной информационной среды увеличивают риски, сопутствующие работе с волатильными финансовыми инструментами.

Эта статья резюмирует трёхлетний опыт создания и оптимизации информационной среды для разработки и использования торговой системы на основе принципов мультитрейдинга, обеспечивающих эффективную работу на нескольких торговых площадках с множеством финансовых инструментов и индикаторов технического анализа тенденций изменения цен [1].

1. ФЕНОМЕН ИНТЕРНЕТ-ТРЕЙДИНГА

В конце XX века появились форекс-брокеры (от англ. Foreign Exchange — «зарубежный обмен») [2], предоставляющие юридическим и физическим лицам интернет-сервисы (торговые платформы) для самостоятельного совершения инвестиционных действий (трейдинга) на международных финансовых рынках с целью извлечения прибыли от изменения курсовой стоимости финансовых инструментов (валютные пары, индексы, металлы, энергоресурсы, сельхозпродукция и пр.). Рыночная цена финансовых инструментов стала зависеть не только от используемых в фундаментальном анализе объективных экономических показателей и политических событий [3], но и от непосредственной активности части человеческой популяции (трейдеров), интересующейся трейдингом. Торговые платформы обеспечили беспрецедентную оперативность открытия, модификации и

закрытия сделок, а также визуализацию изменения цены и её анаморфоз [4] с момента начала обмена данными через интернет.

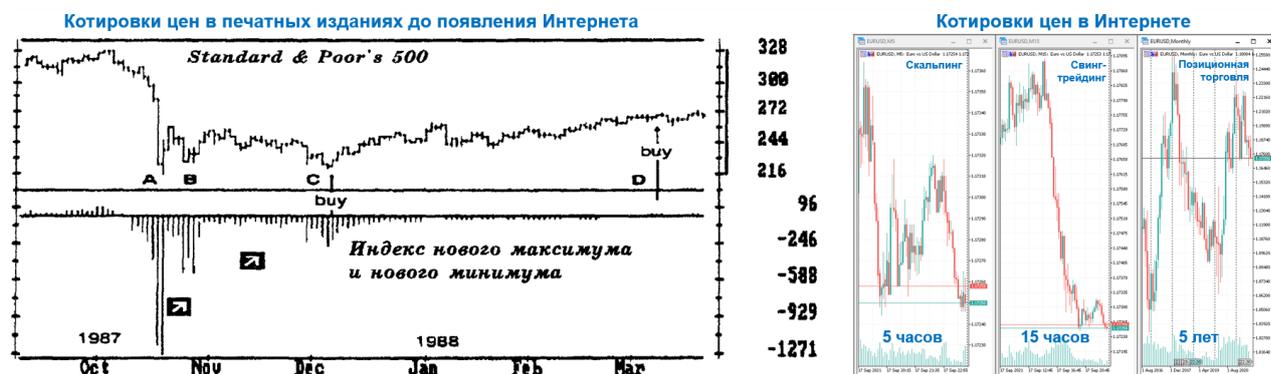


Рис. 1. Котировки цен финансовых инструментов в печатных изданиях (слева) и торговых интернет-платформах (справа). Фрагмент иллюстрации доклада на конференции «Научный сервис в сети Интернет» [1], с дополнениями.

Котировки цен финансовых инструментов до появления интернет-сервисов публиковались в специализированных печатных изданиях. Изменять масштаб визуализации ценовых графиков в печатных изданиях не представлялось возможным, в связи с чем единственной стратегией до появления интернета была позиционная торговля. Интернет-сервисы обеспечили возможность оперативного распространения информации об актуальных котировках цен и масштабирования их графиков, что привело к появлению стратегий свинг-трейдинга и скальпинга для внутридневной торговли (см. Рис. 1).

А. ФОРЕКС-ПОСРЕДНИКИ

Законодательство не разрешает проводить сделки с валютой напрямую. Посредниками между трейдерами и рынком являются форекс-брокеры и дилеры. Торговая платформа обеспечивает оперативное взаимодействие трейдера с интернет-сервером посредника. В течение четверти века эволюции интернет-трейдинга появились десятки торговых платформ [5], используемых многочисленными форекс-посредниками [6]. Оптимальный выбор торговой платформы и связанного с ней форекс-посредника влияет на все аспекты формирования компетенций трейдера.

До появления торговых платформ сделки открывались и закрывались осуществляющим клиринг биржевым брокером (Futures Commission Merchant) по

распоряжениям трейдера, передаваемым через самый быстрый канал связи (голубиная почта, телеграф, телефон). До сих пор солидные брокерские сервисы предоставляют всем аффилированным трейдерам персональных менеджеров, которые по телефонному звонку трейдера или сообщению в онлайн-чате помогут решить проблемы в случае дисфункций торговой платформы, нарушения интернет-связи или затруднений с вводом-выводом средств. Если после регистрации аккаунта трейдера в интернет-сервисе брокера к трейдеру не был прикреплен персональный менеджер с альтернативными электронными каналами оперативной связи (телефон, закрытый чат), то следует продолжить поиски дееспособного брокера.

Для осуществления легальной финансовой деятельности на рынке биржевой брокер должен иметь актуальную лицензию от государственного регулятора в стране, резидентом которой является трейдер. Брокер может не осуществлять регистрацию новых аккаунтов в странах, которые находятся под международными санкциями или определены Группой разработки финансовых мер борьбы с отмыванием денег (Financial Action Task Force on Money Laundering — FATF) как «государства с высоким уровнем риска и не сотрудничающие страны, имеющие стратегические недостатки в сфере противодействия отмыванию преступных доходов и финансированию терроризма (ПОД/ФТ)» [7].

Форекс-брокер информирует трейдера о рыночных котировках через торговую платформу, в которой трейдер оперирует своими личными деньгами от собственного имени. Основным источником дохода брокера – комиссия, которая взимается с каждой сделки. Поэтому брокеры заинтересованы, чтобы трейдеры зарабатывали и сотрудничали с ними как можно дольше. Для привлечения трейдеров брокеры предоставляют возможность использовать в сделках суммы, превышающие депозит трейдеров в 50–1000 раз за счет предоставляемого кредитного плеча (1:50 – 1:1000, соответственно). Брокеры формируют ленты экономических новостей и календари ключевых событий, составляют прогнозы на основе аналитических исследований, проводят семинары и организуют курсы, помогающие начинающим трейдерам.

Форекс-дилер закупает на бирже или у других посредников валюту от своего имени и перепродаёт её трейдерам, самостоятельно формируя в торговой

платформе котировки финансовых инструментов и определяя направление движения денег (биржа, банк, баланс дилингового центра). Таким образом, котировки цен на одинаковые финансовые инструменты у форекс-брокеров и форекс-дилеров могут не совпадать. Форекс-дилер получает доход за счёт изменений цен между покупкой и продажей (спреда). По этой причине спред на торговой платформе форекс-дилера шире, что снижает окупаемость сделок трейдера. Поскольку форекс-дилерам выгодно аккумулировать денежные средства трейдеров и торговать ими на бирже от своего имени, они увеличивают маржинальные требования (минимальный размер средств на торговом счете), предоставляя трейдерам уменьшенное кредитное плечо (1:10 – 1:40).

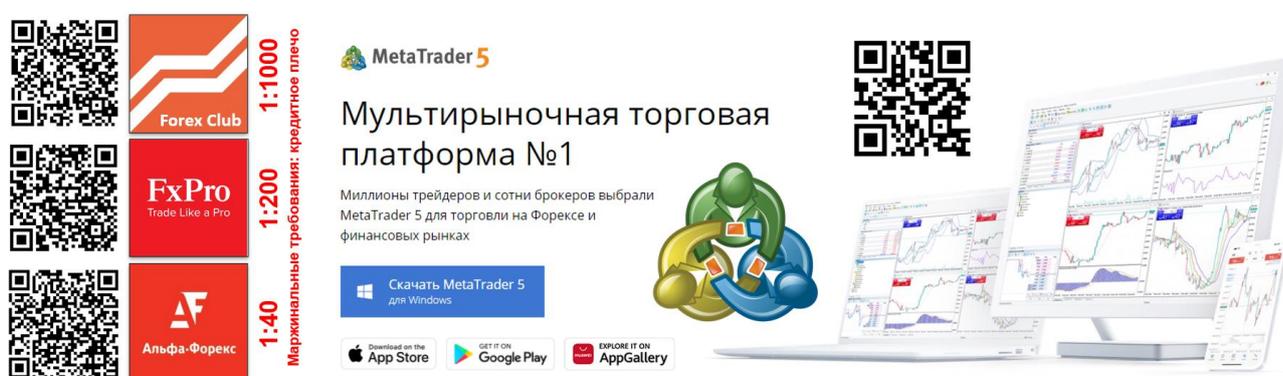


Рис. 2. Форекс-посредники с мультирыночной торговой платформой MetaTrader 5, адаптированной к основным операционным системам и типам пользовательских устройств. Фрагмент иллюстрации доклада на конференции «Научный сервис в сети Интернет» [1].

Первичный выбор форекс-посредников осуществляется трейдером на основе рейтингов [6], региональной доступности, сопоставления функционала торговых платформ и форекс-терминалов, минимального объема сделок (0.01–0.1 лота), спецификаций контрактов и маржинальных требований (см. Рис. 2). На втором этапе определяются оптимальные варианты торговых счетов для реализации различных стратегий.

В. ТОРГОВЫЕ ПЛАТФОРМЫ

Современные торговые форекс-терминалы [5] позволяют открывать, модифицировать и закрывать сделки, оптимизировать технические инструменты и до-

полнять их, использовать существующие торговые стратегии и разрабатывать собственные, транслировать торговые сигналы или принимать их. Первоначально торговые платформы представляли собой исполняемые (exe) приложения, устанавливаемые на персональный компьютер трейдера. Распространение мобильных устройств в начале второго десятилетия XXI века создало условия для появления браузерных веб-терминалов и мобильных приложений, специализированных для смартфонов и планшетов доминирующих операционных систем (см. Рис. 2).

Лицензированные форекс-брокеры и дилеры модифицируют полученную от разработчика торговую платформу посредством добавления символики, привязки к своим серверам, языковой локализации и вариаций настроек торгового инструментария. Посредники-мошенники внедряют в торговые платформы вредоносный код, позволяющий перехватывать управление, отображать ложные котировки финансовых инструментов и бесконтрольно выводить финансовые средства. Вредоносные модификации были обнаружены автором этой статьи в терминалах двух торговых платформ из десятка, апробированного в течение трёх лет: UTIP [8] от брокера vtb-trade.org и Trading Desk Pro 5 [9] от брокера Forex EuroClub. Таким образом, следует после установки новой торговой платформы проверить её файлы коммерческой антивирусной программой с эвристическим анализом и постоянно отслеживать соответствие котировок финансовых инструментов на торговых платформах разных посредников.

Посредники предоставляют трейдерам возможность выбора разных форекс-платформ, к которым при открытии привязываются торговые счета. В дополнение к торговым счетам некоторые торговые площадки [10] практикуют создание виртуальных кошельков, аккумулирующих незадействованные в торговле резервные средства. Смена торговой платформы после открытия счёта невозможна. Действия по вводу, выводу и перераспределению финансовых средств трейдера между различными счетами торговой площадки брокера могут осуществляться в браузерной версии личного кабинета, специальном приложении или непосредственно в торговом терминале. Примером объединения финансово-организационных и торговых функций является современная версия платформы Libertex [11] от брокера Forex Club [12] или FxPro Direct от брокера FxPro [10], при условии использования специального счета FxPro Edge. Современная торговая платформа

(см. Рис. 2) должна обеспечивать работу с одним торговым счётом на форе́кс-терминалах всех типов устройств (компьютер, планшет, смартфон), работающих на основе актуальных операционных систем (Windows, iOS, Android).

Для организации мультитрейдинга [1] необходима торговая платформа, поддерживающая режим многооконной работы с множеством финансовых инструментов и графических индикаторов. Компьютерные и браузерные форе́кс-терминалы позволяют оперативно переключаться между различными счетами торговой площадки только одного брокера, поэтому при формировании информационной среды мультитрейдинга целесообразно устанавливать на компьютер специализированные терминалы всех используемых брокеров, а при использовании веб-терминалов устанавливать отдельный браузер для каждого брокера. Приложения для мобильных устройств позволяют переключаться между торговыми счетами различных брокеров, но этот процесс сопряжен с временной утратой оперативного контроля за состоянием рынка. Адаптированные для планшетов, смартфонов и веб-браузеров форе́кс-терминалы не поддерживают одновременное отображение совокупности окон. При открытии множества вкладок браузера с форе́кс-терминалами в некоторых из них может происходить спонтанное прекращение отображения котировок финансового инструмента вследствие отключения от торгового счёта без какого-либо уведомления.

Утилитарность настройки информационной среды мультитрейдинга определяется функционалом сохранения шаблонов настроек окон и их совокупностей для последующего применения на множестве устройств трейдера с торговыми платформами разных посредников. Эта возможность не реализована в версиях торговых платформ, адаптированных для планшетов и веб-браузеров. Таким образом, основой для аналитики мультитрейдинга могут быть форе́кс-терминалы, оптимизированные для компьютеров. Терминалы для веб-браузеров, смартфонов и планшетов целесообразно использовать для реализации принятых решений по открытию, изменению или закрытию сделок.

В настоящее время наиболее мощным функционалом и надёжностью работы отличается разработанная компанией MetaQuotes [13] торговая платформа MetaTrader 5 (см. Рис. 2), включающая более 80 инструментов технического анализа с возможностью дополнения, 21 таймфрейм вместо 9 традиционных, 6 видов отложенных ордеров и гибкие настройки под любые торговые стратегии.

Важной особенностью MetaTrader 5 является поддержка социального трейдинга [14], обеспечивающего возможность распространять свои торговые сигналы в сообществе трейдеров для копирования с автоматическим исполнением на подключённых торговых счетах. Форекс-дилеры, рекомендуемые трейдерам из Российской Федерации [15, 16], также предоставляют торговые площадки на базе MetaTrader 5. Веским аргументом в пользу выбора посредника трейдером является обеспечение свободного выбора любого варианта форекс-терминала: установка исполняемого файла для компьютера [17], загрузка приложений для мобильных устройств из App Store, Google Play и AppGallery, ссылка на страницу веб-терминала (см. Рис. 2).

С. РЕГУЛИРОВАНИЕ РИСКОВ ТРЕЙДИНГА

Согласно эмпирическому правилу основателя технического анализа Чарльза Доу [18], «90% начинающих трейдеров теряют 90% своего депозита в первые 90 дней торговли».

Для защиты неквалифицированных трейдеров от потенциальных разочарований в Российской Федерации законодательно установлена совокупность условий, требуемых для работы с финансовыми инструментами [19]. Лицензированные форекс-дилеры в Российской Федерации имеют право допускать неквалифицированных инвесторов только к обмену валюты между своими торговыми счетами. Для работы с финансовыми инструментами требуется успешно сдать аттестационный тест (7 вопросов с 4 вариантами ответа каждый).

Основания для получения статуса «Квалифицированный инвестор» [19] : обладание имуществом на сумму от 6 млн рублей (учитываются только деньги на счетах и депозитах в банках, договоры обезличенных металлических счетов и ценные бумаги); присутствие опыта работы в организации, которая совершала сделки с ценными бумагами или производными инструментами (нужно два года опыта, если организация — квалифицированный инвестор, и три года — в других случаях); получение экономического образования в вузе, который мог аттестовывать в сфере профессиональной деятельности на рынке ценных бумаг; наличие аттестата специалиста финансового рынка, аудитора, страхового актуария, CFA, CIIA, FRM; совершение сделок с ценными бумагами или деривативами на сумму от 6 млн рублей за последний год (каждый месяц должно быть не менее одной

сделки, а каждый квартал — в среднем не менее 10 сделок).

Эффективность и соответствующая рискованность трейдинга пропорциональны величине кредитного плеча форекс-посредника (коэффициент, на который умножается сумма средств на торговом депозите трейдера). В соответствии с действующим в Российской Федерации законодательством [19] неквалифицированным инвесторам не рекомендуется пользоваться кредитным плечом, для аттестованных неквалифицированных трейдеров размер кредитного плеча уменьшается при увеличении рисков, а для квалифицированных трейдеров устанавливается на постоянном максимальном уровне. При использовании кредитного плеча соотношение обеспечения трейдера (залога) и размера его обязательств не может быть меньше 1:50 [19], что означает поднятие уровня принудительного закрытия убыточной позиции (stop out). В зависимости от специфики финансового инструмента типичная величина кредитного плеча у российских форекс-дилеров варьирует от 1:20 до 1:40 [15, 16], а у международных брокеров — от 1:200 [10] до 1:1000 [12]. Величина кредитного плеча может уменьшаться при использовании специфических инструментов (энергоресурсов, металлов, акций и индексов) и увеличении сумм сделок. Форекс-посредники могут уменьшать кредитное плечо и увеличивать спред в 2–5 раз на время регламентных работ на торговых платформах в ночные часы и перед закрытием рынков на выходные и праздничные дни. Следует учитывать, что обновление котировок некоторых финансовых инструментов, таких как криптовалюта (BTCUSD), продолжается и в выходные дни.

Для снижения финансовых рисков рекомендуется ограничивать стоимость одной сделки и суммы всех открытых сделок на уровне 2% и 10% от общего количества средств на торговом счёте.

2. САМОРЕАЛИЗАЦИЯ ТРЕЙДЕРА. ПСИХОЛОГИЯ ТРЕЙДИНГА

Трейдинг может являться средством реализации профессиональных талантов, инструментом самофинансирования, источником проблем или великолепной «гимнастикой» для развития аналитических способностей и личностных качеств (концентрация, целеустремлённость, ответственность, самодисциплина). Выбор спекулятивной стратегии и тактики, торговой платформы и финансовых инструментов, поиск работоспособной комбинации индикаторов, их настройка и

интерпретация определяются особенностями фенотипа трейдера (спецификой индивидуальных физиологических и психологических свойств). К примеру, дальтоникам необходимо заменять китайскую цветовую схему трендовой окраски японских свечей (красный/зеленый) иными вариантами, учитывающими особенности восприятия цветов (см. Рис. 3).

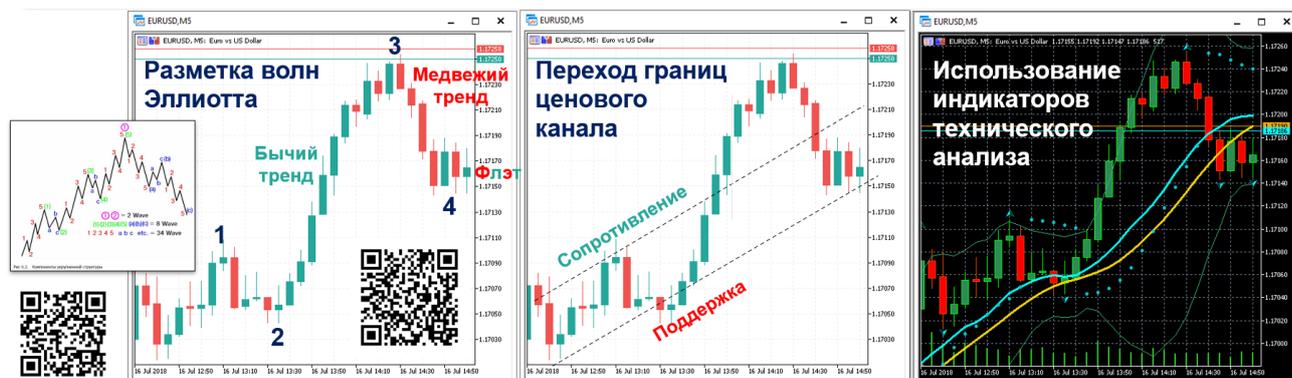


Рис. 3. Варианты прогностического анализа интервального графика котировок цен финансового инструмента EURUSD посредством разметки волн Эллиотта (слева), определения границ ценового канала (в центре), использования индикаторов технического анализа (справа) и представления в форме окрашенных японских свечей. Фрагмент иллюстрации доклада на конференции «Научный сервис в сети Интернет» [1], с дополнениями.

В зависимости от индивидуальных предпочтений трейдеры интуитивно распознают тенденции по совокупности чисел в «биржевом стакане» или графическим тиковым данным, доверительно воспринимают сигналы выхода цены за границу ценового канала; обращают внимание на изменение формы графиков и пересечение сигнальных уровней; реагируют на схождение-расхождение скользящих средних; интерпретируют свечные фигуры и завершённость волновых структур (см. Рис. 3). Одни трейдеры развивают у себя способность к самостоятельному мультитрейдингу (работа с разнообразными индикаторами на нескольких тайм-фреймах каждого из множества финансовых инструментов), а другие программируют и совершенствуют алгоритмы действий торговых роботов (алготрейдинг).

Феноменология трейдинга основана на психологии алчности эмоциональных масс любителей. Ажиотажный спрос и предложение провоцируют импуль-

сивные действия непрофессионального трейдера, приводящие к контрпродуктивным действиям: попытке переждать убытки, отмене страховки (торговля без stop-loss), заключению дополнительных сделок в направлении против старшего тренда (усреднение) или одновременному открытию сделок в противоположных направлениях при неопределённости во флэте (локирование, или «замок»). Практика локирования при отсутствии аналитического инструментария и навыков прогнозности приравнивает любительский трейдинг к азартным играм.

Профессиональный трейдинг несовместим с алчностью и эмоциональностью, прогностической некомпетентностью и поспешностью практических действий. Для выработки профессиональных компетенций форекс-посредники предлагают начинающим трейдерам открывать демо-счета. Однако на демо-счетах по умолчанию устанавливаются крупные размеры торговых депозитов, что не способствует тренировке навыков действий с низкими рисками при минимальном бюджете. Альтернативный вариант последовательной тренировки навыков трейдера: наблюдение за ценовыми графиками и соответствующими индикаторами технического анализа с целью поиска сигнальных корреляций, мысленные открытие и закрытие сделок на реальных счетах, совершенствование торговой системы (набор триггерных индикаторов и условий целесообразности открытия, модификации и закрытия сделок), начало торговли на минимальном депозите с минимальным лотом (не более 2% от депозита), пропорциональное увеличение размера используемых в сделках лотов по мере увеличения количества средств на торговом счете («разгон» депозита). Доходы профессионального трейдера – следствие компетентного прогнозирования изменения цен и грамотного менеджмента средств торгового счёта.

3. СПОСОБЫ ПРОГНОЗИРОВАНИЯ ИЗМЕНЕНИЯ ЦЕН

Фундаментальный анализ используется для прогнозирования цен фондовых инструментов, валюты и биржевых товаров на уровне отдельных компаний, отраслей и экономики стран [3]. Изменение ключевых показателей может привести к разнонаправленному движению цен на разных рынках. Во время выхода новостей с ключевыми показателями начинается кратковременная волатильность рынка, позволяющая заработать на краткосрочных сделках. К примеру, в первую пятницу месяца в 15:30 по московскому времени (MSK) появляются новости об

изменении количества рабочих мест вне сельскохозяйственной отрасли (Nonfarm Payrolls, NFP) и уровня безработицы в США. Таким образом, можно планировать трейдинг в соответствии с расписанием экономических, общественных и политических новостей.

Попытки математического описания закономерностей исторического изменения цен привели к появлению методов технического анализа [4], прогнозирующего тенденцию стабилизации цены (флэт, см. Рис. 3) или её изменения (тренд) в сторону повышения («бычий» тренд, см. Рис. 3) или понижения («медвежий» тренд, см. Рис. 3). Эмпирический анализ изменения цен при чередовании флэтов и трендов привёл к появлению представлений о волновой природе рыночных процессов. Широкое распространение получила теория волн Ральфа Эллиотта [20], в соответствии с которой импульс изменения цены формируется на основе закономерности чисел Фибоначчи из фрактальных элементов, каждый из которых состоит из пяти движений по тренду, перемежающихся с тремя движениями против тренда (см. Рис. 3). Следует учитывать, что в терминологии Эллиотта волной называется однонаправленное импульсное изменение цены, тогда как в физико-математических моделях длина волны определяется как удвоенное расстояние между двумя узлами, при достижении которого точки находятся в одинаковой фазе (полная волна состоит из двух противоположных изменений характеристической величины).

4. КИНЕТИКА ИЗМЕНЕНИЯ ЦЕНЫ ФИНАНСОВОГО ИНСТРУМЕНТА

Согласно современным представлениям, экономические процессы формируются вследствие воздействия разноуровневых осцилляторов (государства, отрасли экономики, компании, физические лица) [21]. Активация осциллятора может быть вызвана новостными событиями или триггерным изменением финансовых потенциалов контрагентов рынка. Для прогноза изменения цены финансовых инструментов составляется модель из суперпозиций отдельных осцилляций. Помимо осцилляций ценовые котировки финансовых инструментов зависят от случайных или спекулятивных флуктуаций [21]. Главная проблема при математическом моделировании осцилляций заключается в определении времени начала действия трендового осциллятора на фоне случайных флуктуаций.

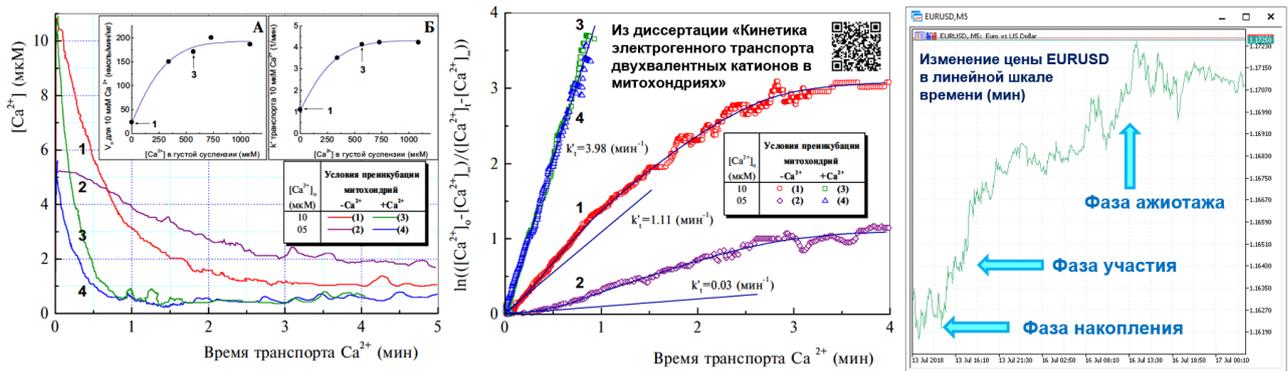


Рис. 4. Аналогия кинетики биологического процесса трансмембранного перераспределения ионов (левая и центральная части [22]) и изменения цены финансового инструмента. Фрагмент иллюстрации доклада на конференции «Научный сервис в сети Интернет» [1], с дополнениями.

Осциллятором может являться коллективное настроение участников рынка. Согласно теории Чарльза Доу [18], формирование первичного тренда начинается без сильных ценовых движений в «фазе накопления» потенциала, когда наиболее проницательные инвесторы начинают действовать вопреки общему мнению (см. Рис. 4, правая часть). Накопление потенциала нового тренда обнаруживается индикаторами технического анализа [4], по сигналам которых действуют активные трейдеры, переводящие тренд в «фазу участия» с сильными изменениями цены. Развитие тренда завершается переходом в «фазу ажиотажа», когда к нему массово присоединяются запоздавшие неквалифицированные трейдеры. В конце «фазы ажиотажа» проницательные инвесторы закрывают прибыльные позиции, что приводит к замедлению ценового движения и переходу во флэт, в течение которого инициируется «фаза накопления» нового тренда. Описанное выше изменение цены во времени аппроксимируется к сигмоидальной кинетике, присущей регулируемым процессам (см. Рис. 4).

Феноменология изменения цены во времени обнаруживает сходство с кинетикой неравновесных химических и биоэнергетических процессов в кооперативных самоорганизующихся системах с дополнительными механизмами регулирования активности [22]. К примеру, при сопоставлении кинетики энергозависимого перераспределения ионов кальция на мембране митохондрий (см. Рис. 4, слева) и изменения котировок EURUSD во времени (см. Рис. 4, справа) наблюдаются аналогии между фазами аллостерической активации транспорта кальция

[22] и накопления торгового потенциала (низкая скорость изменений), активной транслокацией кальция и фазой участия трейдеров, фазой насыщения скорости транспорта ионов и фазой завершения ажиотажной инвестиционной активности участников рынка. Продолжая аналогию, можно уподобить биологические факторы активации транспорта ионов [22] и экономические осцилляторы [21], обеспечивающие активный транспорт ионные каналы [22] и трейдерские торговые терминалы, а также найти сходство в движущей силе перераспределения ионов и денег (электрохимический и финансовый потенциал, соответственно). Обобщение энергетического потенциала в биосистемах происходит на энергопреобразующей мембране [22], а в глобальной экономической среде финансовый потенциал интегрирует единая информационная среда. Общность биологических и экономических процессов перераспределения ресурсов положена в основу гипотезы, представляющей деньги как одну из форм свободной энергии Гиббса, ответственной за выполнение полезной работы в биологических системах [23].

Для определения характеристических кинетических параметров и демаскирования точек инициации латентных лаг-фаз биологических процессов [22] широко используются математические анаморфозы (см. Рис. 4, центр). Разнообразные графические преобразования (анаморфозы) кинетики ценовых изменений и объемов сделок, известные как индикаторы технического анализа [4], могут быть использованы для определения времени начала действия осцилляторов и прогнозирования нелинейных изменений цены во времени.

В результате технологической эволюции торговых платформ появилось множество аналитических индикаторов, указывающих на потенциально важные кинетические параметры изменения цены (направление, ускорение, стабилизация, торможение, амплитуда движения, цены открытия/закрытия сделок и их объёмы). Квалифицированное использование индикаторов технического анализа может увеличивать эффективность трейдинга.

5. ПРИНЦИПЫ МУЛЬТИТРЕЙДИНГА

Информационная среда мультитрейдинга конфигурируется в зависимости от специфики торговых площадок, используемых стратегий, количества торговых счетов и финансовых инструментов.

А. ВАРИАБЕЛЬНАЯ АДАПТИВНОСТЬ ТРЕЙДИНГА

Для позиционной торговли и внутрಿದневного свинг-трейдинга достаточно использовать базовый вариант конфигурации: один компьютер с терминалом MetaTrader 5 [17], в котором настроены профили рабочих областей для обзора состояния всех используемых финансовых инструментов (см. 5.1.5 и 5.5; Рис. 8 и Рис. 13), а также аналитической работы (см. 5.3, Рис. 10 и Рис. 11) и операционных действий (см. 5.4 и Рис. 12) с каждым из них.

Свинг-скальпинг и скальпинг на нескольких торговых счетах с использованием одного финансового инструмента эффективнее осуществлять в информационной среде с распределением аналитических (см. 5.3) и операционных (см. 5.4.1) действий между двумя дисплеями со специфически настроенными терминалами. Поскольку на одном компьютере нельзя одновременно запустить несколько приложений с терминалами одного брокера, для обеспечения аналитической и операционной деятельности выделяются отдельные комплекты компьютеров с дисплеями и устройствами ввода (клавиатура и мышь).

Определение приоритетов в процессе внутридневного трейдинга облегчает третий компьютер, форекс-терминал которого специализирован для Обзора всех используемых финансовых инструментов (см. 5.5 и Рис. 13).

І. МНОЖЕСТВЕННОСТЬ ТОРГОВЫХ СТРАТЕГИЙ

Специфика площадки форекс-посредника и её финансовых инструментов определяет возможность и эффективность использования различных торговых стратегий.

Долгосрочные спекулятивные стратегии [25] с удержанием открытых позиций на срок от нескольких дней до множества лет основываются на фундаментальном анализе [3]. Позиционная торговля (position trading) требует сочетания фундаментального и технического анализов для открытия и закрытия сделок в момент разворота цены на больших таймфреймах – от 1 торгового дня до 2 лет (см. Рис. 1). Специфика позиционной торговли допускает совмещение трейдинга с иными видами занятости.

Внутридневная торговля (intraday trading) осуществляется на основе сигналов индикаторов технического анализа [4]. Стратегия свинг (swing trading) исполь-

зуются на временных интервалах от минут до часов (см. Рис. 1) для открытия/закрытия сделок при изменении направления движения цены во флэте с широким ценовым диапазоном или в продолжительном тренде с перемежающимися импульсными движениями и коррекциями цены.

При использовании стратегии скальпинга (scalping) анализ состояния индикаторов технического анализа дополняется наблюдением за линейными тиковыми графиками (tick chart), отображающими наименьшее изменение котировки финансового инструмента на секундно-минутных таймфреймах (см. Рис. 1) и соответствующими тиковыми объемами сделок.

Трейдинг с использованием внутрисуточных стратегий сложно совмещать с иными видами занятости вследствие необходимости постоянного анализа изменяющейся совокупности индикаторных сигналов.

II. МНОЖЕСТВЕННОСТЬ ТОРГОВЫХ СЧЕТОВ

При открытии торгового счёта у посредника трейдеру предоставляется выбор определения типа счёта – с хеджированием или неттингом. Как правило, форекс-брокеры не допускают изменения типа открытого ранее счёта, тогда как форекс-дилеры предоставляют такую возможность.

Рискованный скальпинг целесообразно практиковать на торговых счетах с хеджированием, допускающим открытие множества независимых сделок в любом направлении (покупка или продажа), которые могут закрываться по-отдельности.

Стратегия свинг-трейдинга с пирамингом предполагает последовательное открытие дополнительных сделок в одном направлении с общим закрытием после получения сигнала о завершении импульса ценового движения от индикаторов технического анализа. Для пираминга оптимизированы торговые счета с неттингом, поддерживающие суммирование открываемых сделок по одному финансовому инструменту с их одномоментным закрытием одним ордером. Для позиционного трейдинга с пирамингом также удобно использовать торговый счёт с неттингом.

III. МНОЖЕСТВЕННОСТЬ ТОРГОВЫХ ПЛОЩАДОК

Для обеспечения надёжности трейдерской деятельности целесообразно дублировать работу с одним финансовым инструментом на однотипных торговых платформах нескольких посредников. Такой мультитрейдинг обеспечивает синхронное формирование буферных депозитов в процессе компаративного анализа работы торговых платформ разных форекс-посредников на случай прекращения деятельности одного или нескольких из них вследствие технических или организационных проблем (санкции, отзыв лицензии и т. п.).

Платформа для скальпинга должна обеспечивать минимальное проскальзывание (временной интервал между отправкой команды трейдером через форекс-терминал и реальным исполнением ордера на сервере брокера). Критически важными условиями для осуществления скальпинга является высокая волатильность (изменчивость во времени) цены финансового инструмента, небольшой спред (разница между максимальной ценой продажи (ask) и минимальной ценой покупки (bid)), быстрая реакция веб-сервера на изменение рыночных котировок и действия трейдера (см. Рис. 5), а также низкая комиссия брокера за открытие сделки. Следует принимать во внимание, что спред увеличивается в нерабочее время и уменьшается при открытии торговых сессий, каждой из которых соответствует свой набор финансовых инструментов.

Выбор торговой площадки обуславливается региональной доступностью посредника, условиями контрактов (см. Рис. 2) и соответствием приоритетной торговой стратегии трейдера. Надёжный способ определения оптимальной торговой площадки для скальпинга – сопоставление в реальном времени тиковых графиков наиболее ликвидной валютной пары EURUSD в окнах торговых платформ разных брокеров. На Рис. 5 видно, что величина спреда и динамика котировок на торговых платформах Forex Club (справа) и Альфа-Форекс (в центре) создают благоприятные условия для скальпинга, позволяя совершить пару прибыльных сделок в течение 2 минут, а на платформе FxPro (слева) такая возможность не появляется.

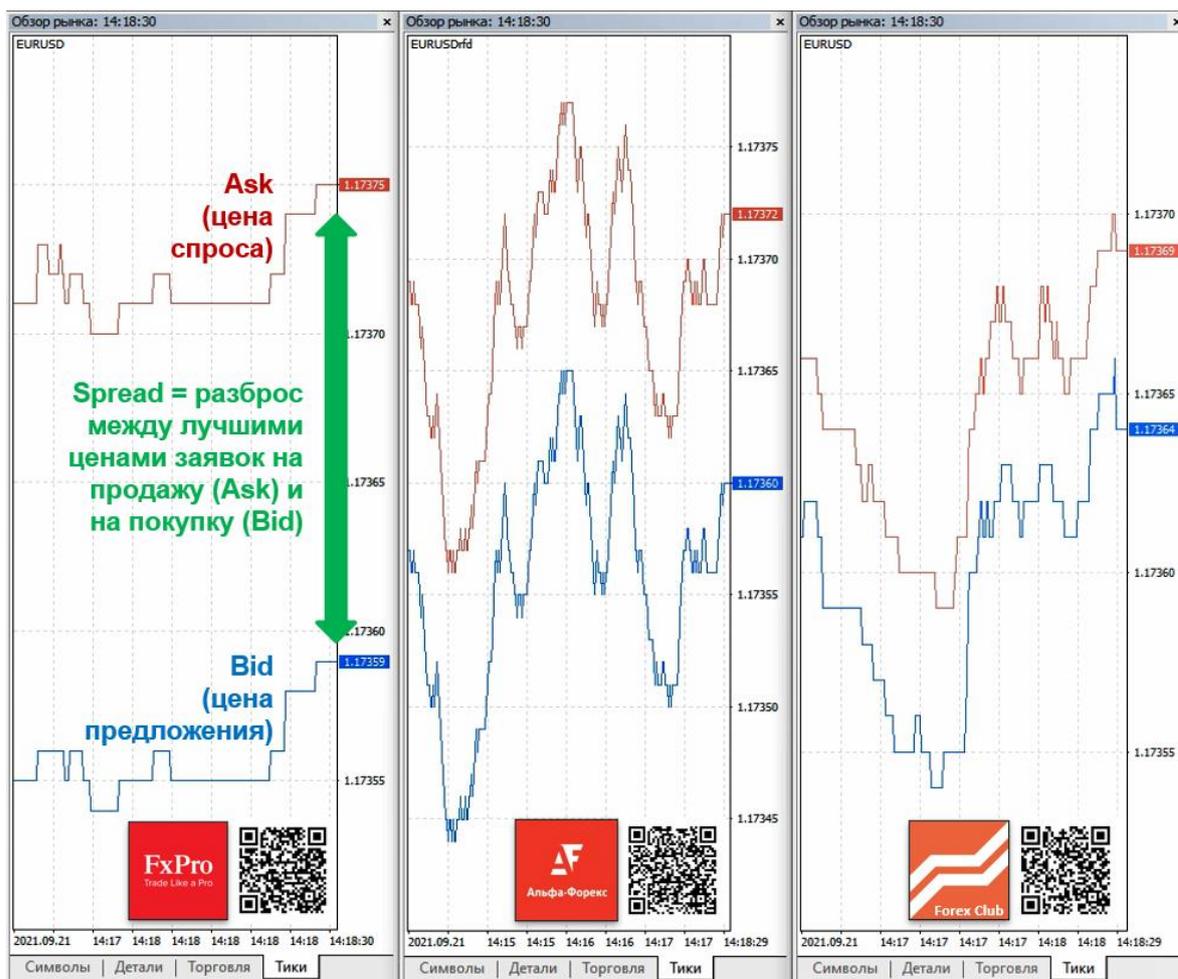


Рис. 5. Синхронные тиковые графики финансового инструмента EURUSD в компьютерных форекс-терминалах торговой платформы MT5 посредников FxPro [10], Альфа-Форекс [15] и Forex Club [17]. Фрагмент иллюстрации доклада на конференции «Научный сервис в сети Интернет» [1].

Свинг-скальпинг целесообразно организовать с использованием разных торговых счетов, один из которых применяется для скальпинга, а другой – для свинг-трейдинга, при котором сделки открываются в точках разворота трендов с продолжительностью от получаса до полусуток. Примером оптимальной платформы для свинг-скальпинга и свинг-трейдинга является MetaTrader 5 у компаний Forex Club [17] и FxPro [10], соответственно. При осуществлении сделок с небольшими лотами (единица купли-продажи финансового инструмента) эффективно использовать посредников с относительной платой, которая зависит от объема торговой позиции.

Позиционный трейдинг эффективнее осуществлять у посредников с фиксированной платой за проведение операций с любым лотом и оптимальными условиями свопа (комиссия за перенос позиций через ночь). При позиционной торговле следует принимать во внимание, что в результате обновления брокерами форекс-терминалов их специфика может существенно изменяться с течением времени.

IV. МНОЖЕСТВЕННОСТЬ ИНДИКАТОРОВ

Форекс-терминалы MetaTrader 5 позволяют использовать множество шаблонов аналитических окон и профилей рабочих областей из нескольких окон. Инсталляционные комплекты шаблонов и профилей komponуются форекс-посредниками (см. Рис. 6).

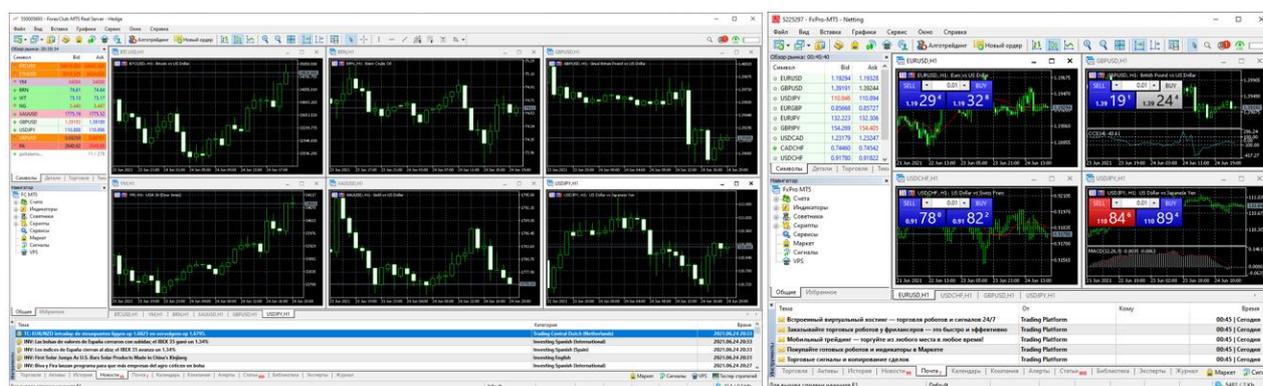


Рис. 6. Базовые профили рабочей области компьютерных форекс-терминалов торговой платформы MetaTrader5 посредников Forex Club (слева [17]) FxPro (справа [10]). Фрагмент иллюстрации доклада на конференции «Научный сервис в сети Интернет» [1].

Опыт показывает, что базовый набор индикаторов неинформативен и требует доработки в соответствии с торговой системой (при наличии) и индивидуальными предпочтениями трейдера. При формировании и оптимизации информационной среды мультитрейдинга целесообразно заменять комплект шаблонов и профилей актуальным вариантом, сохраняя все устаревшие шаблоны и профили в архивах. Для надёжности работы рекомендуется устанавливать форекс-терминалы всех брокеров на системный диск (несъёмный носитель) с учётом выделения по 5 Гб для хранения исторических данных с котировками финансовых инструментов в расчёте на каждый терминал. Данные терминалов располагаются в

папках с буквенно-цифровыми названиями из 32 символов в директории `c:\Users\ИМЯ-ПОЛЬЗОВАТЕЛЯ\AppData\Roaming\MetaQuotes\Terminal\`. Внутренняя структура папок с данными терминалов стандартна: шаблоны настройки одного Аналитического окна терминала и профили рабочей области с множеством настроенных окон находятся внутри папок `...\MQL5\Profiles\Templates\` и `...\MQL5\Profiles\Charts\`, соответственно.

Рекомендуется сохранять каждую новую комбинацию индикаторов в отдельном шаблоне Аналитического окна (см. Рис. 7). Шаблоны можно подключать к однотипным терминалам разных посредников, использовать с любыми финансовыми инструментами и распространять в сообществе трейдеров. Сохранение шаблона в папке `...\MQL5\Profiles\Templates\` происходит по команде, доступной из контекстного меню, вызываемого нажатием правой кнопки мыши в рабочем пространстве Аналитического окна. Для упорядочивания использования потенциально распространяемых шаблонов Аналитических окон целесообразно именовать их по маске `FIOYYYYMMDD`, заменяя `FIO` на инициалы автора шаблона, а `YYYY`, `MM` и `DD` – на год, месяц и день сохранения модификации. Профили рабочих областей сохраняются в папке `...\MQL5\Profiles\Charts\` по команде меню «Файл > Профили > Сохранить как» с наименованиями на основе названий используемых шаблонов Аналитических окон со специфическими приставками для использования в Аналитическом и Обзорном дисплеях. Приставки для названий шаблонов профилей рабочей области Аналитического дисплея целесообразно устанавливать по маске `AI-NAME_`, где `AI` обозначает «Анализ одного Инструмента», а `NAME` заменяется символьной аббревиатурой финансового инструмента, принятой у форекс-посредника. Следует иметь в виду, что наименования одного финансового инструмента у разных форекс-посредников могут отличаться. К примеру, котировки золота по отношению к американскому доллару в терминалах `Forex Club` [17] и `FxPro` [10] обозначаются как `XAUUSD` и `GOLD` соответственно.

При подключении шаблона профиля рабочей области различие наименований финансовых инструментов приводит к невозможности автоматического отображения котировок, что исправляется перетаскиванием мышью символьного обозначения соответствующего финансового инструмента в поле Аналитического

окна из списка, доступного по команде «Вид > Обзор рынка > Символы». Для адекватного отображения содержимого Аналитических окон при подключении шаблона рабочей области к терминалу необходимо присутствие всех задействованных шаблонов Аналитических окон в папке ... \MQL5\Profiles\Templates\.



Рис. 7. Эволюция состава и представления содержимого Аналитического окна с ценовым графиком (сверху) и комбинациями технических индикаторов компьютерного терминала MetaTrader 5 от Forex Club [17]. В верхней части окон ценовых графиков указана дата разработки шаблона в формате YYYYMMDD. Все варианты Аналитического окна отображают синхронные котировки инструмента EURUSD на таймфрейме M15.

Для прогнозирования моментов разворота цен во флэтах в пределах одного ценового диапазона используются индикаторы из группы осцилляторов [4]. Однако при переходе от флэта к тренду сигналы осцилляторных индикаторов теряют значимость. Трендовые индикаторы учитывают кинетику изменения ценового диапазона, в котором нижняя граница определяется уровнем поддержки (потенциал покупки достаточен для остановки падения цен), а верхняя – уровнем сопротивления (потенциал продажи достаточен для остановки роста цен). Середину трендового канала обозначают индикаторы из подгруппы скользящих средних. Индикаторы увеличения объема торгов подтверждают развитие тренда. Следует учитывать, что сигналы трендовых индикаторов утрачивают актуальность при переходе к флэту.

Эффективность позиционной торговли и внутридневного свинг-трейдинга

повышается посредством синхронного анализа сигналов индикаторов разного типа, некоторые из которых не сочетаются с ценовым графиком или друг с другом. По этой причине в окно ценового графика финансового инструмента добавляется несколько окон индикаторов. Оптимальные комбинации индикаторов (см. 5.2) и их настройка (см. 5.7.1) определяются на основании эмпирического опыта и корректируются в соответствии с индивидуальными предпочтениями трейдера и действующими эргономическими нормативами [24].

V. МНОЖЕСТВЕННОСТЬ ФИНАНСОВЫХ ИНСТРУМЕНТОВ

Увеличение спреда, переход из тренда во флэт и замедление изменения цены являются основаниями для смены приоритетно используемого во внутриведневной торговле финансового инструмента. Набор доступных финансовых инструментов определяется спецификой торгового терминала и лицензией форекс-посредника.

Международные форекс-брокеры предоставляют возможность инвестировать средства в сотни разнообразных финансовых инструментов: валюты и криптовалюты, металлы, энергоресурсы (нефть и газ), сельхозтовары, акции и биржевые индексы. Региональные форекс-дилеры, как правило, ограничивают набор финансовых инструментов десятком популярных валютных пар.

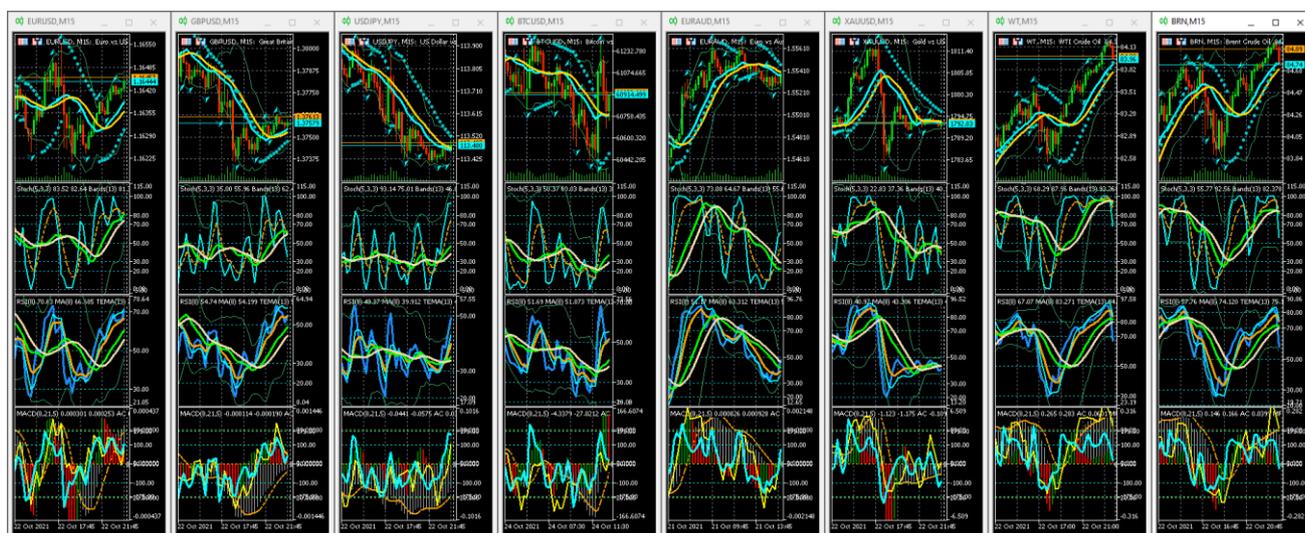




Рис. 8. Фрагменты экранных снимков рабочей области терминала MetaTrader 5 брокера Forex Club [17] на Обзорном дисплее с аналитическими окнами 15-минутных таймфреймов для двух наборов из 8 финансовых инструментов по состоянию на 24.10.2021. Инструменты шаблона профиля O1-M15_FOK20210812: EURUSD, GBPUSD, USDJPY, BTCUSD, EURAUD, XAUUSD, WT, BRN. Инструменты шаблона профиля O2-M15_FOK20210812: SUGAR, WHEAT, COFFEE, COCOA, SOYBEAN, HD, NG, USDRUB.

На начальном этапе формирования торговой системы целесообразно определить основной и дополнительный наборы финансовых инструментов, используя как критерии минимальную величину спреда (см. Рис. 8) и совпадение времени наибольшей волатильности цены инструмента с рабочим временем региона пребывания трейдера. Наибольшая волатильность цены инструмента наблюдается во время открытия торговых сессий в соответствии с региональным распределением источников валют и сырьевых ресурсов (время MSK): Тихоокеанская (00:00–09:00), Азиатская (03:00–12:00), Европейская (10:00–18:00) и Американская (16:00 до 23:00). В набор финансовых инструментов полезно включить национальную валюту страны пребывания трейдера (см. Рис. 8, USDRUB в наборе O2-M15_FOK20210812).

В. ЦЕЛЕВАЯ СПЕЦИАЛИЗАЦИЯ ПРОГНОСТИЧЕСКИХ ИНДИКАТОРОВ

Универсальных торговых индикаторов и их комбинаций не существует. Простейшая тактика действий – найти тренд и присоединиться к нему – оказывается контрпродуктивной в момент окончания тренда. Прибыльная торговля во флэте

на разворотах цены по осцилляторным индикаторам может нивелироваться убытками при появлении и развитии любого тренда. Таким образом, торговую систему следует разрабатывать, комбинируя индикаторы различного типа в одном Аналитическом окне.

Аналитическое окно в рабочей области терминала создаётся при выборе финансового инструмента из меню «Файл > Новый график > ...». Новое Аналитическое окно содержит только ценовой график. Замена графика финансового инструмента осуществляется перетаскиванием мышью в поле окна буквенного обозначения финансового инструмента из списка во вкладке «Символы», появляющейся в левой части рабочего пространства при вызове из меню «Вид > Обзор рынка» (Ctrl+M).

Свойства ценового графика редактируются во вкладках всплывающего окна, вызываемого при выборе пункта меню «Графики > Свойства» (F8). На вкладке «Общие» рекомендуется установить автопрокрутку и интервальный тип графика «Японские свечи», в которых элемент отображает диапазон изменения цены в течение заданного периода. Толстое «тело» свечи символизирует преобладающие цены открытия и закрытия сделок, а границы тонких теней сверху и снизу тела указывают на абсолютные максимумы и минимумы цены за период. Цвет свечи соответствует тренду периода: белый или зелёный (в зависимости от цветовой схемы) обозначает «бычий» тренд, а черный или красный – «медвежий» тренд. Соотношения длин тела и теней свечей, а также комбинации из 2–3 свечей различной формы (фигуры) помогают прогнозировать направление движения цен и определять волатильность (рыночный разброс цен спроса и предложения). На вкладке «Показывать» полезно активировать отображение всех элементов, кроме «Показывать реальные объемы» и «Показывать описания объектов». Для наглядной визуализации спреда важно включить пункты «Показывать линию Bid» и «Показывать линию Ask». Если данные торговой истории начинают мешать восприятию графиков, то следует деактивировать пункт «Показывать торговую историю».

С ценовым графиком можно совмещать некоторые индикаторы технического анализа (см. Рис. 7, верхнее окно всех шаблонов). Добавление индикаторов

к активному окну осуществляется выбором пунктов меню «Вставка > Индикаторы > ...» или перетаскиванием в поле Аналитического окна элементов древовидной структуры «Индикаторы», вызываемой из меню «Вид > Навигатор». Набор индикаторов определяется используемой торговой стратегией и индивидуальными предпочтениями трейдера. Оптимизация набора индикаторов (дополнение, сокращения и настройка параметров) происходит в процессе создания индивидуальной торговой системы.

Некоторые индикаторы технического анализа совместимы с ценовым графиком Аналитического окна. К примеру (см. Рис. 7, верхнее окно шаблона 20210503), представление об актуальных сигналах разворота движения цены при скальпинге и свинг-трейдинге может формироваться при нахождении корреляции меток Фрактального индикатора (Fractals, стрелки, цвет – Aqua) с разворотами Фрактальной адаптивной скользящей средней (Fractal Adaptive Moving Average, линия, цвет – Aqua). Для определения границ диапазона вариаций цены полезны Ленты Боллинджера (Bollinger bands, цвет – MediumSeaGreen во всех шаблонах), рассчитываемые на основе стандартного отклонения от собственной Скользящей средней (Moving Average, цвет – Aqua в шаблоне от 20210812).

Пересечение средней линии Боллинджера и производной от её данных дополнительной Скользящей средней (Moving Average, цвет – Gold в шаблоне от 20210812) при соответствующей настройке может быть подтверждающим сигналом смены основного тренда. Цвет линий парных индикаторов подбирается таким образом, чтобы при бычьем тренде сверху оказывалась линия из синей части спектра, а при медвежьем – из красной (см. Рис. 7, шаблоны от 20181127–20210812, а также 5.7.1). Расширение и сжатие пространства между Лентами Боллинджера сигнализируют о начале и завершении импульсного движения цены, соответственно. Подключение к ценовому графику Параболического индикатора (Parabolic Stop and Reverse, или Parabolic SAR, цвет – DarkTurquoise в шаблоне от 20210812) позволяет узнавать рекомендуемые уровни закрытия убыточных сделок (Stop Loss).

Индикаторы, несовместимые с ценовым графиком (собственные значения ординат), при добавлении к Аналитическому окну вызывают его разделение на верхнее окно Главного графика цены и нижнее Окно индикатора 1 (см. Рис. 7, вто-

рое окно сверху). В Окне индикатора 1 можно разместить множество индикаторов с подходящими параметрами. Аналитическое окно может содержать несколько окон индикаторов, которые группируются по совместимости, специфике (трендовые, осцилляторы, индикаторы объёмов) и целевому назначению (упреждающие, основные, подтверждающие).

Главная проблема простых индикаторов технического анализа – запаздывание торговых сигналов относительно времени появления значимых изменений цены. На Рис. 9 видно, что сигнал МА (пересечение двух скользящих средних) в Аналитическом окне таймфрейма М4 формируется через 30 минут после начала медвежьего тренда (совокупность красных японских свечей в конце ценового графика).

Анализ истории изменения цен позволяет предположить, что кинетика флэта может быть описана гармоническими осцилляциями цены при постоянстве среднего уровня во времени, а кинетика тренда соответствует осцилляциям цены с направленным по тренду смещением среднего уровня во времени и дисбалансом амплитуд отклонения цены от скользящей средней в направлении тренда. Таким образом, при развитии тренда амплитуда изменений цены в серии осцилляций прогрессивно увеличивается с одновременным уменьшением амплитуды контртрендовых изменений. Наклон скользящей средней отражает отклонение движущих сил (активности продавцов и покупателей) от равновесия.

В результате экспериментов с последовательным подключением трендовых индикаторов к осцилляторным для сглаживания несущественных колебаний цены и минимизации количества ложных сигналов при определённой совокупности настроек было обнаружено демаскирование границ фаз изменения цены. Оказалось, что все фазы тренда можно идентифицировать по наклону и местонахождению авторского Комплексного Индикатора Тенденций (CIT, Complex Indicator of Tendencies), который формируется из нескольких трендовых индикаторов, последовательно подключаемых к данным осцилляторного технического индикатора изменения цены. В результате компаративных экспериментов было установлено, что наибольшей прогностической силой отличаются комбинации Скользящих средних Лент Боллинджера (графики цвета Lime в окнах Упреждающих и Основных индикаторов на Рис. 9), применённых к данным осцилляторов (Stoch и RSI)

со Скользящими средними, применёнными к данным Лент Боллинджера (графики цвета Wheat в окнах Упреждающих и Основных индикаторов на Рис. 9). В соответствии с названиями используемых осцилляторов (Stoch и RSI), Комплексные Индикаторы Тенденций получили наименования CITS и CTR, соответственно.

Нахождение линий индикатора CИТ ниже или выше уровня 50% в совокупности с отрицательным или положительным наклоном указывает на «медвежий» или «бычий» тренд, соответственно. Наклонное параллельное расположение линий CITS старшего таймфрейма свидетельствует об ускорении тренда в соответствующем наклону направлении (см. Рис. 9). Горизонтальное совпадение линий CИТ указывает на стабильность тенденций: при уровнях ниже или выше 50 % прогнозируются «медвежий» или «бычий» тренды, соответственно, а мелкие осцилляции вблизи уровня 50% обозначают флэт.

Не исключено, что эволюция принципа построения комплексных индикаторов тенденций на основе комбинаций осцилляторных и трендовых индикаторов окажется продуктивной для создания аналитических систем прогнозирования кинетики изменения ключевых параметров самоорганизующихся кооперативных систем в биологии, кибернетике, экономике и социальной сфере.

Таким образом, для определения трендов осцилляций целесообразно использовать Комплексные Индикаторы Тенденций, создаваемые последовательным применением трендовых индикаторов к осцилляторным индикаторам изменения цены. В зависимости от специфики осциллятора Комплексные Индикаторы Тенденций могут формировать упреждающие или основные торговые сигналы (см. 5.2.1 и 5.2.2.). При распределении компонентов в вертикальном стеке Аналитического окна целесообразно под ценовым графиком последовательно размещать окна индикаторов с упреждающими, основными и подтверждающими сигналами (см. Рис. 7, эволюция к шаблону 20210812 и Рис. 9).

I. УПРЕЖДАЮЩИЕ ИНДИКАТОРЫ

Комплексный Индикатор Тенденций CITS, компонуемый на основе данных Стохастического осциллятора (Stoch), формирует упреждающие сигналы о предстоящей смене тренда и позволяет заблаговременно прогнозировать направление выхода из консолидационного флэта или контртрендовой коррекции (см. Рис. 9). Угол наклона линий CITS в момент пересечения уровня 50 коррелирует с

накопленным во время консолидации финансовым потенциалом и соответствующей силой предстоящего трендового изменения цен.

Известно, что при анализе осцилляций наиболее проблематичным является определение начала первой волны. Было установлено, что Первая волна Эллиотта демаскируется индикатором CITS, сформированным на основе Стохастического осциллятора, и помечается крестообразным пересечением линий Стохастического осциллятора с линиями CITS, инвертирующимися в маргинальной зоне (см. Рис. 9).

Главная функция упреждающих индикаторов – сигнализация о начале и завершении осцилляционного изменения цены на краях тренда, а также прогнозирование амплитуды осцилляций.

II. ОСНОВНЫЕ ИНДИКАТОРЫ

Комплексный Индикатор Тенденций CITR, komponуемый на основе данных осциллятора RSI, формирует торговый сигнал во время инверсии составляющих его скользящих средних в момент начала трансформации осцилляций по начинающемуся тренду. Главная функция основных индикаторов – подтверждение появления тренда и сигнализация момента начала быстрых высокоамплитудных изменений цены. Сигналом предстоящего ценового скачка является пересечение уровня 50% Скользящей средней, применённой к данным индикатора RSI (см. Рис. 9, оранжевые графики в окнах Основных индикаторов). Расположение и наклон этой Скользящей средней коррелируют с характером предстоящих осцилляторных изменений цены в диапазоне актуального тренда. Разворот вышеупомянутой Скользящей средней у линий CITR после их пересечения в момент инверсии указывает начало 3-й волны Эллиотта [20].

III. ПОДТВЕРЖДАЮЩИЕ ИНДИКАТОРЫ

Для определения вероятности окончания или продолжения тренда используются подтверждающие индикаторы. К примеру, значительное отклонение индикатора MACD (серые столбики на Рис. 9 в окне Подтверждающих индикаторов) от среднего уровня предупреждает об опасности открытия сделки по сигналам Упреждающих и Основных индикаторов в момент пересечения уровня 50 линиями CITS и CITR. И наоборот, совпадение линий индикатора Money Flow Index и

Commodity Channel Index (см. MFI+CCI на Рис. 9) подтверждает целесообразность открытия сделок с максимально доступным лотом в соответствии с сигналами Упреждающих и Основных индикаторов до момента пересечения уровня 50 Скользящей средней, подключенной к индикатору RSI (оранжевые графики в окне Основных индикаторов на Рис. 9).

С. СТРАТЕГИЧЕСКАЯ ВИЗУАЛИЗАЦИЯ ФРАКТАЛЬНОГО КОНТИНУУМА ИСТОРИИ ИЗМЕНЕНИЯ ЦЕНЫ ФИНАНСОВОГО ИНСТРУМЕНТА

Технический анализ с использованием разных временных масштабов (таймфреймов) позволяет демаскировать фрактальную сущность кинетики изменения цен (периодическое самоподобие структур). Зарождение и завершение тренда лучше прослеживаются на младших таймфреймах (с меньшим периодом) а продолжение – на старших (с большим периодом).

При сопоставлении синхронного изменения цены одного финансового инструмента на разных таймфреймах идентифицируются тренды, актуальные для используемой торговой стратегии [25]; обнаруживаются вероятные уровни поддержки и сопротивления; прогнозируется амплитуда изменения цены и демаскируется фрактальная природа рыночных процессов.

Современные компьютерные торговые терминалы позволяют организовать одновременное наблюдение за изменениями цен финансового инструмента в двух десятках окон с разными таймфреймами. Определение набора таймфреймов, способствующего прогностической аналитике, осуществляется на основании компаративных экспериментальных наблюдений. Отдельная задача – оптимизация размещения окон с разными таймфреймами в пределах одной торговой платформы.

Использование нескольких Аналитических окон одного финансового инструмента с разными таймфреймами для увеличения эффективности прогнозов в отношении направления изменения цены было предложено А. Элдером [26]. В соответствии с торговой системой «трёх экранов» А. Элдера [26] индикаторы Аналитического окна с крупным масштабом времени используются для стратегического определения безопасного открытия сделок в направлении актуального крупного тренда (см. Рис. 9). Тактическое определение ожидаемой амплитуды изменения цены осуществляется по сигналам индикаторов Аналитического окна со

средним масштабом времени. Для определения моментов открытия и закрытия сделок при начале и окончании трендового изменения цены используются интерпретации состояния индикаторов Аналитического окна с наименьшим масштабом времени. Таким образом, базовый комплект рабочей области терминала целесообразно формировать из трёх Аналитических окон с разными таймфреймами (см. Рис. 9). Актуальная совокупность таймфреймов трёх Аналитических окон может смещаться в соответствии с кинетикой изменения цен. Определение перспективных тенденций помогает своевременно переназначать основной торговый таймфрейм в пределах совокупности окон Аналитического дисплея.

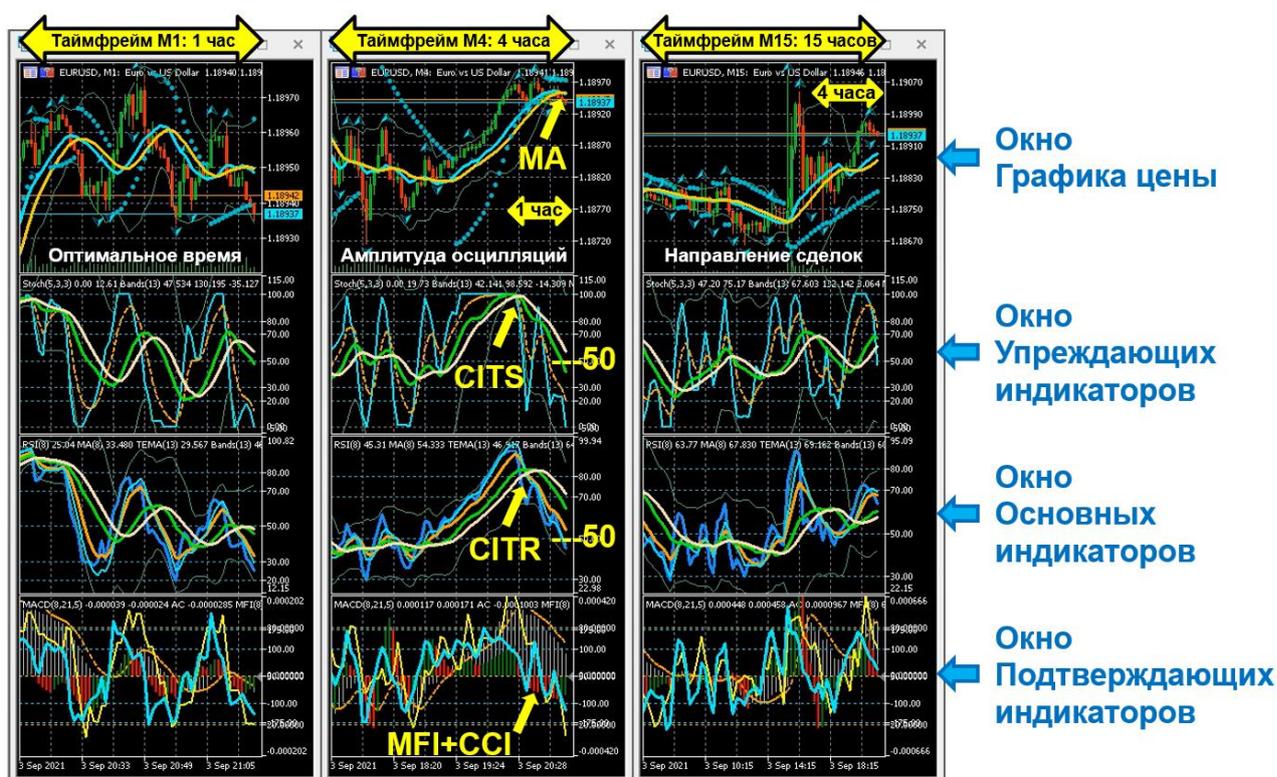


Рис. 9. Базовый комплект рабочей области Аналитического дисплея из трёх Аналитических окон с масштаб интервальных графиков 1 минута (таймфрейм M1), 4 минуты (таймфрейм M4) и 15 минут (таймфрейм M15). Стрелками указаны торговые сигналы: MA – скользящие средние; CITS – упреждающий Комплексный Индикатор Тенденций на основе осциллятора Stoch; CITR – основной Комплексный Индикатор Тенденций на основе осциллятора RSI, MFI+CCI – подтверждающие индикаторы Money Flow Index и Commodity Channel Index. Фрагмент иллюстрации доклада на конференции «Научный сервис в сети»

Интернет» [1], с дополнениями.

Были проведены эксперименты по сопоставлению эффективности прогнозирования изменения цены посредством базового комплекта из трёх Аналитических окон с фиксированной разницей таймфреймов в 3, 4, 5, 8 раз или с прогрессивным увеличением таймфреймов, определяемых последовательным умножением масштаба времени на числа Фибоначчи (1 минута, 3 минуты, 15 минут, 2 часа, 1 день, 1 месяц). Оказалось, что интерпретация совокупности состояния всех индикаторов рабочей области терминала облегчается при фиксированной 4-кратной разнице таймфреймов Аналитических окон комплекта, при которой правая четверть Аналитического окна со старшим таймфреймом отображает всё содержимое Аналитического окна с ближайшим младшим таймфреймом (см. Рис. 9 и Рис. 10).

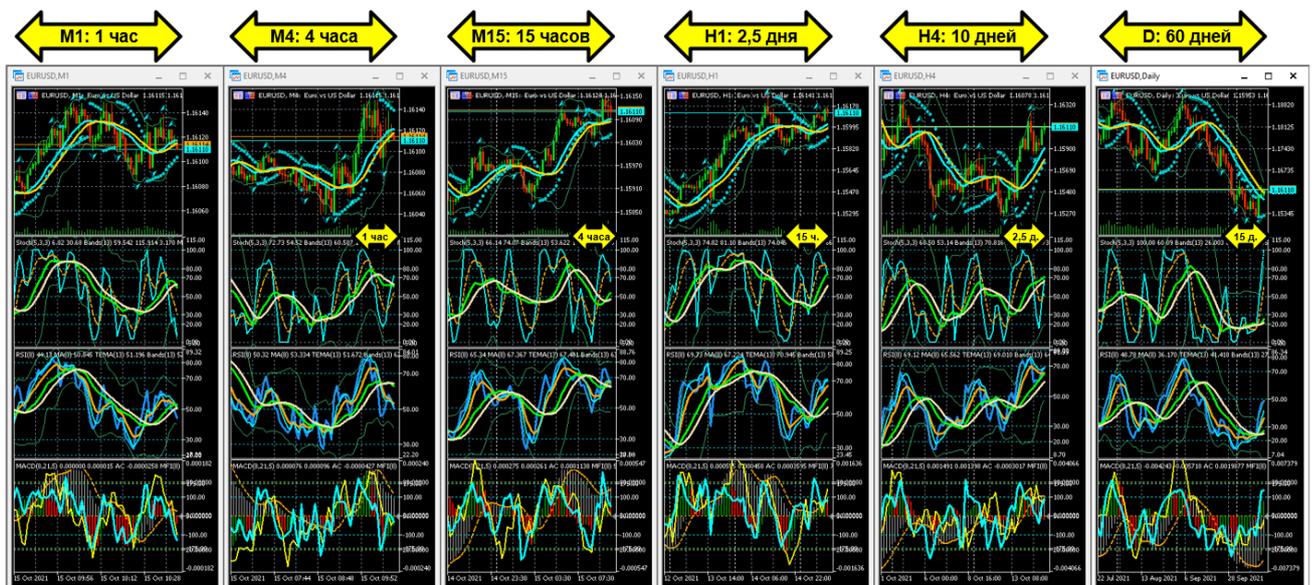


Рис. 10. Комплект окон Аналитического дисплея с 6 таймфреймами для свинг-трейдинга (шаблон 20210812). Фрагмент иллюстрации доклада на конференции «Научный сервис в сети Интернет» [1], с дополнениями.

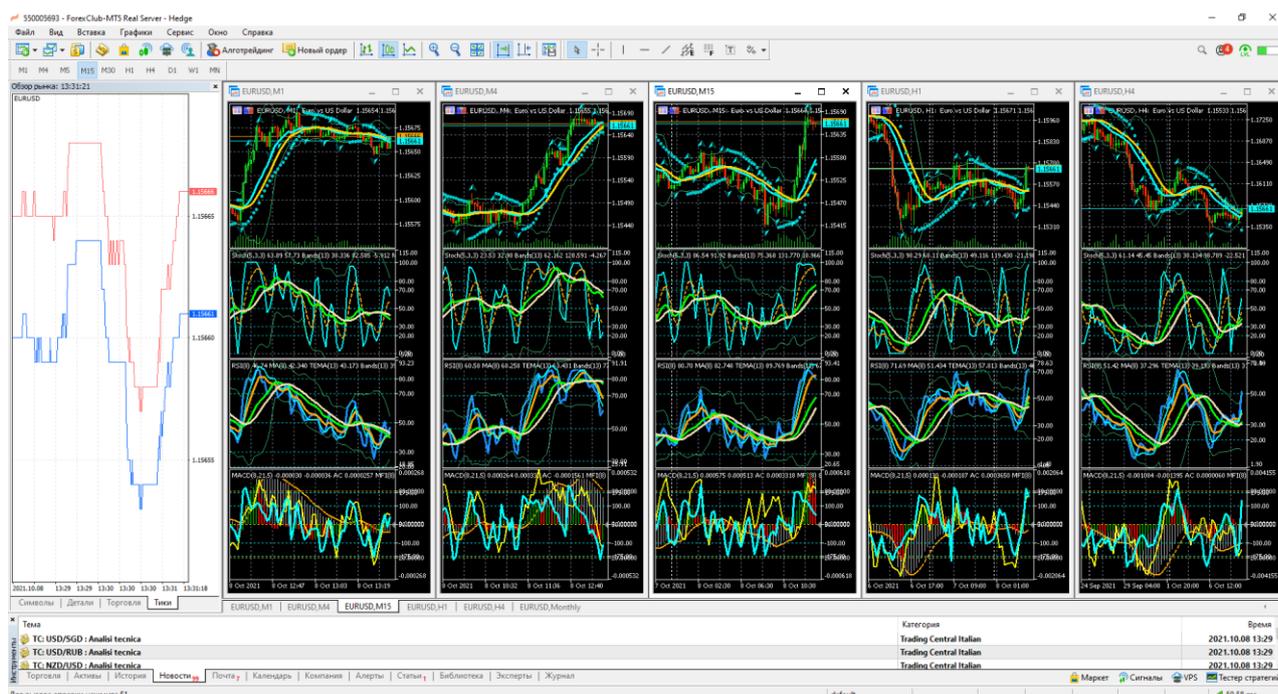


Рис. 11. Экранный снимок терминала MetaTrader 5 от Forex Club [17] на Аналитическом дисплее инструмента EURUSD с тиковым графиком и Аналитическими окнами 5 таймфреймов для свинг-скальпинга (шаблон 20210812).

При размещении комплекта Аналитических окон в рабочем пространстве терминала на дисплее с альбомной ориентацией (ширина:высота = 16:9) по команде «Окно/Вертикально» разномасштабные графики цен и индикаторов располагаются горизонтальными рядами, что упрощает сопоставление их содержания. Поскольку скорость и амплитуда изменения цены финансовых инструментов варьируют во времени, актуальные таймфреймы Аналитических окон для оперативного, тактического и стратегического планирования могут изменяться. По этой причине рабочее пространство целесообразно комплектовать расширенным набором Аналитических окон с 5–7 различными таймфреймами, в пределах которых может смещаться актуальная совокупность «трёх экранов А. Элдера». При комплектации рабочего пространства терминала шестью Аналитическими окнами устанавливается оптимальное для восприятия аспектное соотношение графиков цен и соответствующих окон индикаторов (см. Рис. 10).

Для позиционной торговли комплект Аналитического дисплея продуктивно компоновать из окон с таймфреймами 15 минут, 1 час, 4 часа, 1 день, 1 неделя, 1

месяц. Компоновка рабочего пространства терминала Аналитического дисплея для свинг-трейдинга осуществляется с использованием окон с меньшими таймфреймами: 1 минута, 4 минуты, 15 минут, 1 час, 4 часа, 1 день (см. Рис. 10). При такой конфигурации Аналитического дисплея содержимое Аналитических окон разного масштаба отображает набор соответствующих промежутков времени, используемых при всех стратегиях (скальпинг, свинг-трейдинг, позиционная торговля): M1 – 1 час, M4 – 4 часа, M15 – 15 часов, H1 – 2,5 суток, H4 – 10 суток, D – 3 месяца.

В терминале для скальпинга Аналитические окна целесообразно предварить тиковым графиком окна «Обзор рынка» (см. Рис. 11), вызываемого из меню «Вид > Обзор рынка > Тики» (Ctrl+M).

D. ОПЕРАТИВНОЕ ВЫПОЛНЕНИЕ ДЕЙСТВИЙ НА МНОЖЕСТВЕ ТОРГОВЫХ СЧЕТОВ

При использовании различных стратегий на одной платформе трейдер может открыть несколько торговых счетов и использовать их одновременно, что требует соответствующей реорганизации рабочего пространства. Работа с множеством торговых платформ повышает общую эффективность трейдинга при наступлении форс-мажорных обстоятельств в период глобального системного кризиса (блокировка доменных зон, IP-адресов брокеров, санкции в отношении участников движения денежных средств, ограничение набора доступных финансовых инструментов, неисправность торговых платформ, нарушение работы сети Интернет).

I. КОНФИГУРАЦИЯ РАБОЧЕГО СТОЛА ОПЕРАЦИОННОГО ДИСПЛЕЯ

Особенность трейдинга с использованием краткосрочных сделок (скальпинг и свинг-скальпинг) – необходимость организации быстрого исполнения ордеров без отрыва от анализа сигналов технических индикаторов Аналитического дисплея и принятия соответствующих решений. Для реализации принятых решений в форекс-терминале вызывается Окно управления ордером (Order), которое заслоняет 20% рабочего пространства, препятствуя продолжению анализа всего комплекса индикаторов. В компьютерном терминале горизонтальный размер

Окна Ордера можно сократить вдвое, убрав отображение тикового графика из левой половины посредством двойного клика левой кнопки мыши. Вертикальный размер Окна Ордера изменить нельзя. В терминале веб-браузера Окно ордера не масштабируется.

Процесс вызова Окна Ордера требует переключения внимания и совершения дополнительных действий. В компьютерном терминале новый ордер открывается пунктом меню «Сервис > Новый ордер», кнопкой «Новый ордер» или клавишей F9, а в веб-терминале – только кнопкой «Новый ордер». Окно редактирования открытого ордера вызывается двойным кликом строки ордера на вкладке «Торговля» нижнего окна терминала, открываемого пунктом меню «Вид > Инструменты». Высоту окна «Инструменты» можно варьировать в соответствии с количеством одновременно открытых сделок, сокращая площадь всех Аналитических окон терминала, что приводит к снижению различимости сигналов индикаторов. Включение режима «Торговля в один клик» не рекомендуется вследствие высокого риска возникновения ошибок при работе с ордерами.

Переключение терминала между торговыми счетами осуществляется в окне, которое вызывается из пункта меню «Файл > Подключиться к торговому счёту» и закрывает 10% площади рабочего пространства терминала. Таким образом, при использовании торговых стратегий, требующих быстрых действий, совмещение в одном терминале аналитической и операционной деятельности контрпродуктивно.

Информационная среда мультитрейдинга с несколькими торговыми счетами разных брокеров может быть оптимизирована посредством выделения отдельных комплектов оборудования для осуществления аналитической и операционной деятельности. Операционный дисплей должен обеспечивать быстрое выполнение ордеров на всех торговых счетах без переключения между ними. С этой целью в пространстве рабочего стола мозаично размещаются окна веб-терминалов с заблаговременно вызванными Окнами ордеров, в которых осуществляются открытие, модификация или закрытие сделок (см. Рис. 12).

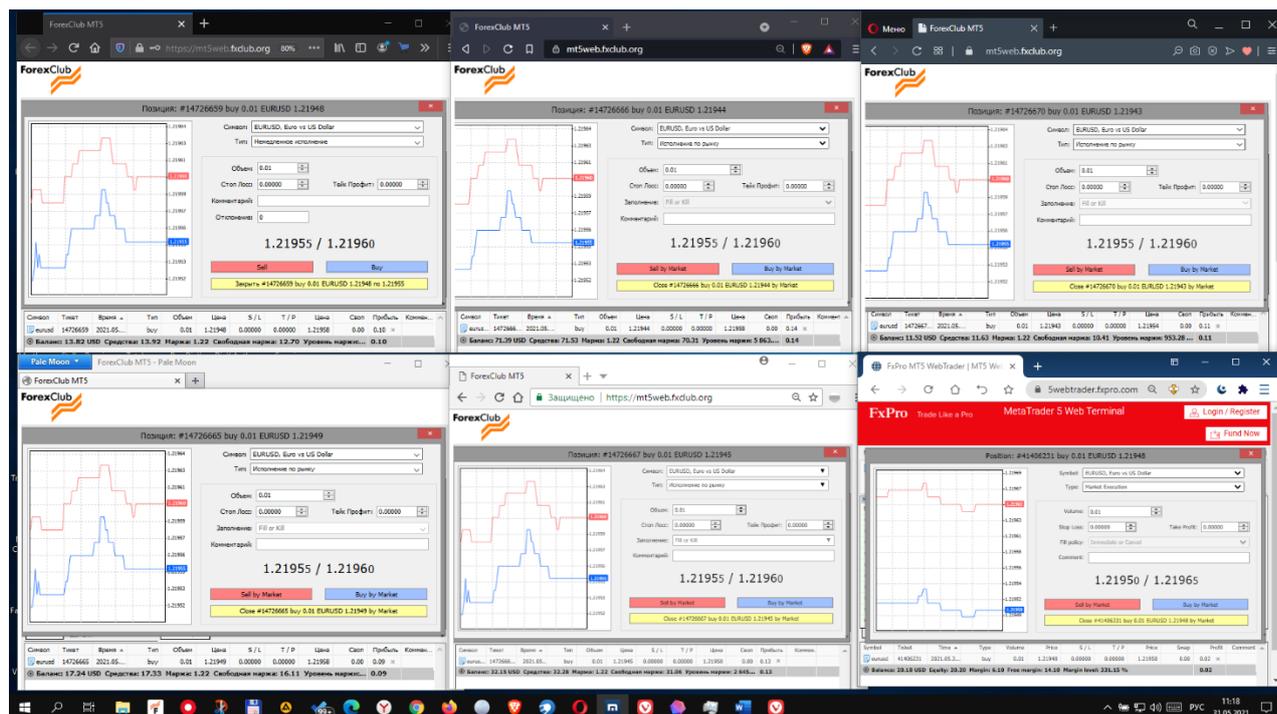


Рис. 12. Экранный снимок рабочего стола Операционного дисплея с открытыми Окнами ордеров в 6 веб-терминалах MetaTrader 5, подключённых к нескольким торговым счетам брокеров Forex Club [17] и FxPro [10].

Приёмы конфигурирования содержимого рабочего стола для операционной деятельности мультитрейдинга сводятся к оптимизации распределения окон с веб-терминалами в соответствии с приоритетными стратегиями для торговых счетов на различных брокерских площадках. Если у каждого брокера трейдер работает только с одним торговым счётом, то рабочий стол Операционного дисплея можно конфигурировать с использованием компьютерных веб-терминалов, специализированных брокерами. При необходимости работы с несколькими счетами у одного брокера предпочтительными элементами Операционного дисплея являются окна веб-браузеров с онлайн-терминалами, подключёнными к соответствующим счетам.

В одном браузере можно открыть множество вкладок с веб-терминалами, подключить к ним разные торговые счета с одним финансовым инструментом, разобщить вкладки по отдельным окнам и распределить окна в пространстве рабочего стола для оперативного управления ордерами. Поскольку смена финансовых инструментов в работающих веб-терминалах вкладок одного браузера может приводить к прекращению отображения содержимого окон с графиками цены,

рекомендуется открывать веб-терминал каждого используемого счёта в отдельном браузере, отличном от используемого в системе по умолчанию. Такая организация информационной среды Операционного дисплея увеличивает отказоустойчивость и позволяет быстро конфигурировать рабочий стол, поскольку предварительно настроенные окна каждого браузера открываются в predeterminedённых для них местах (общая конфигурация рабочего стола не подлежит сохранению).

Дополнительные браузеры целесообразно установить в систему до начала организации содержимого рабочего стола Операционного дисплея, отключить в них блокировку трекеров и настроить автоматическое открытие веб-терминалов соответствующих брокеров при запуске. Для работы с веб-терминалами подходят браузеры, способные непрерывно обеспечивать круглосуточную работу одной вкладки, экономно расходуя ресурсы центрального процессора (1–5%) и операционной памяти (50–200Мб): Brave (<https://brave.com>), Comodo Dragon (<https://browser.comodo.com>), Epic (<https://www.epicbrowser.com>), Maxthon (<https://www.maxthon.com>), Osiris (<https://browseosiris.com>), Pale Moon (<https://www.palemoon.org>), Sputnik (<https://sputnik-lab.com>), Vivaldi (<https://vivaldi.com>).

Веб-браузеры с активными голосовыми помощниками (Yandex – <https://browser.yandex.ru>) для трейдинга использовать не рекомендуется. Для размещения браузеров с веб-терминалами на Операционном дисплее рекомендуется выделять отдельный рабочий стол. При выборе инструментария Операционного дисплея следует учитывать, что некоторые браузеры искажают отображение ценовых графиков Opera (<https://www.opera.com>) или клонируют контуры своего окна между открытыми рабочими столами (Avant Browser – <http://www.avantbrowser.com>).

II. КОНФИГУРАЦИЯ ВЕБ-ТЕРМИНАЛОВ ОПЕРАЦИОННОГО ДИСПЛЕЯ

Веб-терминал открывается в браузере по URL-адресу, предоставляемому брокером. После окончания загрузки страницы веб-терминала следует подключиться к торговому счёту и актуализовать содержимое Графика цены перетаскиванием в его поле символа используемого финансового инструмента из окна «Обзор рынка». Для других операций График цены в веб-терминалах Операционного

дисплея не используется. Главным функциональным элементом веб-терминала Операционного дисплея является Окно ордера, а вспомогательным компонентом служит окно «Инструменты» со списком открытых ордеров. Для конфигурирования окна браузера с веб-терминалом нажатием кнопки «Новый ордер» следует вызвать Окно ордера и переместить его в крайнюю доступную позицию у левого верхнего угла веб-терминала.

В веб-терминале нижняя граница Графика цены устанавливается на уровне нижней границы Окна ордера таким образом, чтобы находящееся в крайнем верхнем положении Окно ордера полностью закрывало окна «Обзор рынка» и «График цены», оставляя видимым заголовки столбцов окна «Инструменты». Настройка высоты окна браузера зависит от разрешения Операционного дисплея и осуществляется таким образом, чтобы по вертикали на рабочем столе помещались окна двух соседних браузеров.

Высота веб-браузера минимизируется посредством перемещения нижней браузера вверх таким образом, чтобы в окне «Инструменты» были видны строки с информацией о балансе счёта и параметрах открытого ордера. При необходимости целесообразно уменьшать масштаб отображения страницы веб-терминала в браузере. К примеру, для дисплея с разрешением FullHD в альбомной ориентации (1920x1080) используется масштаб 70–80%, в зависимости от специфики браузера. Ширина окон браузеров минимизируется до правой рамки Окна Ордеров. Такая конфигурация браузеров позволяет мозаично разместить на рабочем столе Операционного дисплея шесть веб-терминалов (см. Рис. 5), подключённых к разным счетам одного или нескольких брокеров на любых комбинациях финансовых инструментов.

Веб-терминалы брокеров распределяются по браузерам Операционного дисплея в соответствии с оптимальными стратегиями для соответствующих торговых счетов. В результате компаративных экспериментов было установлено, что эффективность операционной деятельности повышается в результате размещения веб-терминалов для скальпинга, свинг-скальпинга и свинг-трейдинга в левой, центральной и правой частях Операционного дисплея, соответственно (см. 5.7.3).

Е. ТАКТИЧЕСКИЙ ВЫБОР ПЕРСПЕКТИВНОГО ФИНАНСОВОГО ИНСТРУМЕНТА

Продуктивность трейдинга снижается при формировании длительного флэта с низкоамплитудным диапазоном медленных осцилляций цены финансового инструмента между уровнями поддержки и сопротивления. В такой ситуации необходим выбор альтернативного финансового инструмента посредством сопоставления аналитических данных в совокупности окон множества финансовых инструментов. Критерии выбора и финансовых инструментов и вариантов их отображения варьируются в зависимости от стратегии трейдинга.

Для выбора наиболее перспективного финансового инструмента используется Обзорный дисплей, отображающий компьютерный терминал с набором Аналитических окон разных финансовых инструментов с одинаковым таймфреймом (см. Рис. 8). В результате экспериментов по конфигурированию профиля рабочей области Обзорного дисплея было установлено, что оптимальное количество Аналитических окон Обзорного дисплея при «альбомной» и «книжной» ориентациях монитора равно 6 и 8 соответственно. Наилучшее отображение 8 Аналитических окон разных финансовых инструментов обеспечивает монитор с классическим соотношением сторон 4:3 (см. Рис. 13). В информационной среде мультитрейдинга [1] Обзорный дисплей обслуживается отдельным компьютером и располагается справа от стека из Аналитического и Операционного дисплеев (см. 5.7.4 и Рис. 14).

На этапе формирования торговой системы целесообразно использовать совокупность профилей рабочей области терминала с расширенным набором таймфреймов, который используется для установления наиболее информативного масштаба времени Аналитических окон, предопределяемого приоритетной стратегией. Для свинг-трейдинга удобен таймфрейм M15, обеспечивающий обзор всего внутрисуточного торгового интервала времени (см. Рис. 8 и Рис. 13).



Рис. 13. Экранный снимок рабочей области терминала MetaTrader 5 брокера Forex Club [17] на Обзорном дисплее с аналитическими окнами 15-минутных таймфреймов для 8 финансовых инструментов (EURUSD, GBPUSD, USDJPY, BTCUSD, EURAUD, XAUUSD, WT, BRN).

Г. ИНТЕГРАТИВНОСТЬ ИНФОРМАЦИОННОЙ СРЕДЫ ТРЕЙДИНГА

На начальных стадиях разработки торговой системы мультитрейдинга (настройка компонентов Аналитического дисплея) достаточно использовать один комплект оборудования (компьютер, монитор, клавиатура и мышь).

Одновременно с мониторингом на Аналитическом дисплее котировок активно используемого финансового инструмента можно осуществлять тактический выбор перспективного финансового инструмента из множества доступных, задействовав дополнительный комплект оборудования с Обзорным дисплеем. Эффективная работа с множеством торговых счетов возможна при использовании третьего комплекта оборудования с Операционным дисплеем.

Интегративная работа Обзорного, Аналитического и Операционного дисплеев (см. Рис. 14) может поддерживаться безвентиляторными микрокомпьютерами [27] с соответствующим размером оперативной памяти (2, 4 и 8 Gb), индивидуальными дисплеями размером 21"–28" и общим комплектом «клавиатура и мышь», поддерживающим технологию Logitech Flow. Эта технология обеспечивает перемещение курсора мыши в сопряженное инфопространство соседнего компьютера при пересечении границы экрана с соответствующим автоматическим переключением фокуса клавиатуры.

Клавиатура Logitech Craft снабжена аппаратным колесом, которое удобно использовать для скроллинга содержимого аналитических окон и видеозаписей дисплеев мультитрейдинга. Для надёжности рекомендуется подключать мышь и клавиатуру к компьютерам не посредством Bluetooth, а при помощи специализированных USB-передатчиков Logitech Unifying.

Было экспериментально подтверждено, что безвентиляторные микрокомпьютеры способны поддерживать работу Аналитического, Операционного и Обзорного дисплеев системы мультитрейдинга одновременно с видеозахватом экрана в круглосуточном режиме 5 дней в неделю в течение трёх лет и более. Отказоустойчивость аппаратных компонентов системы мультитрейдинга обеспечивается источниками бесперебойного питания и качественной сетевой инфраструктурой.



Рис. 14. Рабочее место для мультитрейдинга с использованием трёх дисплеев: Аналитического (слева сверху), Операционного (слева снизу) и Обзорного (справа). В Аналитических окнах применён шаблон 20210812. Иллюстрация из доклада на конференции «Научный сервис в сети Интернет» [1], с изменениями.

Инфосреда мультитрейдинга может быть модифицирована для работы в мобильных условиях: стационарные дисплеи и комплект «клавиатура+мышь» заменяются тремя сенсорными экранами с разрешением не менее 1920x1080 или одним экраном с матричным HDMI-переключателем визуальных сигналов (пример: Navseker ZY-HM22) и тремя беспроводными микроклавиатурами (пример: IPazzPort 19s). Связь с мобильными точками доступа рекомендуется осуществлять посредством внешних антенн с достаточным коэффициентом усиления сигнала (пример: WD-1900AC RTL8814).

При возникновении дисфункций компонентов в процессе работы отказо-

устойчивость интегративной инфосреды мультитрейдинга обеспечивается единством набора фореक्स-терминалов, веб-браузеров, а также шаблонов Аналитических окон и профилей рабочей области, которые синхронизируются после каждой модернизации между всеми задействованными комплектами оборудования через системы хранения в облаке или локальное RAID-хранилище.

Г. ЭСТЕТИЧНАЯ ЭРГОНОМИКА ИНФОРМАЦИОННОЙ СРЕДЫ ТРЕЙДИНГА

Работа профессионального трейдера требует большой концентрации внимания. В составе одного Аналитического окна (см. Рис. 7, шаблон 20210812) скомпонованы график цены и 16 индикаторов технического анализа. Аналитический и Обзорный дисплеи содержат 6 и 8 Аналитических окон, в которых находятся 102 и 136 графических объектов соответственно. Таким образом, общее число объектов непрерывного мониторинга в информационной среде мультитрейдинга достигает 238. Поскольку современный СанПин 1.2.3685-21 [24] рекомендует ограничивать число объектов наблюдения (от 6 до 10) и плотность сигналов (от 76 до 175 в среднем за час работы), при формировании инфосреды мультитрейдинга эргономике уделяется особое внимание.

I. ЭРГОНОМИЧЕСКАЯ ОПТИМИЗАЦИЯ АНАЛИТИЧЕСКИХ ОКОН

Как показывает опыт, базовая настройка компонентов фореक्स-терминалов не обеспечивает информативность и визуальный комфорт графически представляемых данных (см. Рис. 6). По этой причине трейдеру следует определить необходимый избыточный набор компонентов Аналитического дисплея и взаимно адаптировать их оформление (цвет, толщина, тип линии) с целью маскирования несущественных и выделения важных сигнальных элементов.

При настройке шаблона Аналитического окна (см. Рис. 7) решаются две задачи: совмещение ценового графика с множеством производных от него индикаторов и оптимизация графического представления компонентов окна для длительного ожидания формирования торговых сигналов в совокупности индикаторов технического анализа. Принцип целевой группировки (см. 5.2) индикаторов с применением Комплексных Индикаторов Тенденций позволяет уменьшить втрое количество отслеживаемых графиков в каждом Аналитическом окне (см. Рис. 9).

Эксперименты по оптимизации цветового оформления ценового графика (вкладка «Цвета») завершились выбором цветовой схемы «Green on Black» и определением цветов элементов на основе китайского стиля: Фон – Black; Текст – White; Сетка – LightSlateGray; Бар вверх – Lime; Бар вниз – OrangeRed; Бычья свеча – SeaGreen; Медвежья свеча – Red; Линия – Lime; Объёмы – LimeGreen; Линия Bid – Aqua; Линия Ask – Orange; Линия последней цены – 0,192,0; Стоп-уровни – Red (см. Рис. 7, верхнее окно всех шаблонов). Такая настройка облегчает концентрацию внимания трейдера на важных элементах ценового графика и снижает нагрузку на зрение при длительной работе.

Оформление индикаторов подбирается в соответствии со способом их подключения (к значениям цены или данным других индикаторов), значимостью и относительной позицией. В соответствии с физической природой (синяя часть спектра соответствует большим энергиям, а красная – меньшим) цвет линий парных индикаторов устанавливается таким образом, чтобы при бычьем тренде сверху оказывалась линия из синей части спектра, а при медвежьем – из красной (см. Рис. 7, шаблоны 20181127–20210812).

Для концентрации внимания трейдера на существенных сигналах первичные индикаторы с «шумными» данными маскируются посредством визуализации тонкими или пунктирными линиями, а вторичные индикаторы со значимыми сигналами выделяются толстыми линиями (см. Рис. 7, шаблон 20210812). Во избежание когнитивной перегрузки рекомендуется создавать внутри Аналитического окна не более трёх Окон индикаторов (см. Рис. 7, три нижних окна в шаблонах 20200518–20210812 и Рис. 9).

Неизбыточная достаточность обозреваемого промежутка времени обеспечивается ограничением ширины Аналитического окна на уровне 60 японских свечей (см. Рис. 7, шаблоны 20210503, 20210812), что обеспечивает лучшую визуализацию актуальной истории изменения цены на всех таймфреймах. В результате компаративных экспериментов было установлено, что оптимальное соотношение ширины и высоты для Главного графика равно 4:3, а для каждого из трёх Окон индикаторов – 16:9 (см. Рис. 7, шаблоны 20210503, 20210812).

II. ЭРГОНОМИЧЕСКАЯ ОПТИМИЗАЦИЯ ОБЗОРНОГО ДИСПЛЕЯ

Эксперименты по оптимизации конфигурирования Обзорного дисплея завершились выбором мозаичного размещения восьми Аналитических окон на экране с книжной ориентацией и классическим аспектным соотношением сторон 3:4 (ширина : высота). При соотношении сторон экрана 9:16 (ширина : высота) целесообразно уменьшать количество Аналитических окон Обзорного дисплея до шести. Определение набора финансовых инструментов, подлежащих мониторингу, осуществляется трейдером на основе индивидуальных предпочтений. Мозаика Аналитических окон Обзорного дисплея организуется таким образом, чтобы инструменты с положительной корреляцией изменения цен располагались в одном столбце, а сверху находился инструмент, в индикаторных окнах которого формируются упреждающие торговые сигналы. К примеру, целесообразно размещать EURUSD над GBPUSD и WT над BRN (см. Рис. 13).

III. ЭРГОНОМИЧЕСКАЯ ОПТИМИЗАЦИЯ ОПЕРАЦИОННОГО ДИСПЛЕЯ

Рабочий стол Операционного дисплея рекомендуется конфигурировать таким образом, чтобы в верхнем ряду и нижнем ряду оказывались веб-браузеры с тёмной и светлой темой оформления, соответственно. Известно, что приоритетным для внимания среднестатистического человека является верхний левый угол монитора. По этой причине веб-терминал для скальпинга целесообразно размещать в левом верхнем углу Операционного дисплея (см. Рис. 12).

Если количество торговых счетов превышает полдюжины, для размещения дополнительных веб-терминалов целесообразно создавать отдельный рабочий стол. Следует учитывать, что при выключении компьютера рабочие столы закрываются, но в системе сохраняются относительное расположение всех веб-браузеров и их привязка к определённому рабочему столу.

IV. ЭРГОНОМИЧЕСКАЯ ОПТИМИЗАЦИЯ ОБОРУДОВАНИЯ ДЛЯ МУЛЬТИТРЕЙДИНГА

Эффективность аналитической работы, принятия ответственных решений и их выполнения зависит от поддержания высокого уровня сосредоточения трейдера. Шум вентиляторов нескольких компьютеров увеличивает нагрузку на нерв-

ную систему. По этой причине для формирования информационной среды мультитрейдинга целесообразно использовать бесшумные безвентиляторные микро- и миникомпьютеры [27], такие как ACEPC и MeLE.

Было определено, что размещение Аналитического дисплея над Операционным (см. Рис. 14) способствует минимизации нагрузки на шейный отдел позвоночника трейдера при долговременной работе (оба дисплея устанавливаются на кронштейн в альбомной ориентации). Для рационального размещения двух мониторов и двух компьютеров удобно использовать наклонно-поворотный кронштейн (стандарт VESA), предназначенный для четырёх мониторов (примеры: KRON D422E со струбцинным креплением и KRON D421FS для установки на столешницу).

В результате экспериментов с различными моделями мониторов (разрешение, тип матрицы, размер) было установлено, что длительное наблюдение за состоянием индикаторов Аналитического дисплея и работа с ордерами Операционного дисплея облегчаются при размере экранов 27"–28" с разрешением не менее 1920x1080. Оптимальное разрешение Обзорного дисплея – 1200x1600 при размере 22"–24". Наименьшее утомление глаз обеспечивают мониторы с матрицами MVA. Итоговая конфигурация мониторов стационарной системы мультитрейдинга: Аналитический дисплей iiyama X2783HSU-B3, Операционный дисплей iiyama X2888HS-B2, Обзорный дисплей NEC MultiSync 2190UXp.

Эргономическая оптимизация взаимного размещения трёх мониторов системы мультитрейдинга завершилась выбором варианта, в котором Обзорный дисплей в книжной (вертикальной) ориентации располагался справа от укрепленных друг над другом Аналитического и Операционного дисплеев в альбомной (горизонтальной) ориентации (см. Рис. 14). При выборе модели Обзорного дисплея предпочтителен вариант с подставкой, рассчитанной на регулировку по высоте, наклон и поворот монитора.

6. ИНФОРМАЦИОННЫЕ СЕРВИСЫ ТРЕЙДИНГА

Для осуществления интернет-трейдинга достаточен двухсторонний обмен данными между биржевым посредником и трейдером. Следует принимать во

внимание, что использование дополнительных интернет-сервисов может существенно разнообразить деятельность трейдера и увеличить продуктивность его времяпрепровождения.

А. СТАТИСТИКА ИНВЕСТИЦИОННОЙ ДЕЯТЕЛЬНОСТИ

Анализ эффективности торговой стратегии, тактики и операционной деятельности существенно облегчается при автоматизации сбора статистических данных активности трейдера. Непосредственно после установки терминала MetaTrader 5 по умолчанию происходит вызов веб-браузера с переходом по адресу страницы сервиса сообщества пользователей приложений MQL5 [14]. Регистрация трейдера в сервисе позволяет подключить к нему терминалы всех посредников и мобильные устройства трейдера для обмена торговыми данными, сигналами и новостями.

В. СОЦИАЛЬНЫЙ ТРЕЙДИНГ. ТОРГОВЫЕ СИГНАЛЫ

Трейдер имеет возможность публиковать данные своей работы с торговым счётом в форме торгового сигнала (см. Рис. 15) посредством сервиса, предоставляемого разработчиками форекс-терминала MetaTrader 5 [14]. Аналитическая система сервиса позволяет оценить стиль работы трейдера и эффективность его торговой системы. Любой зарегистрированный пользователь сообщества MQL5 имеет возможность подключить торговый сигнал трейдера к своему терминалу и активировать автоматическое копирование сделок.

Следует учитывать, что эффективность копирования сделок зависит от используемой стратегии: при скальпинге система может не успевать за действиями трейдера. По этой причине целесообразно не использовать разные стратегии на одном торговом счёте и называть торговые счета в соответствии с применяемой стратегией (см. Рис. 15). При отсутствии торговой активности в течение месяца сигнал отключается и помещается в архив, откуда его можно извлечь по инициативе трейдера после совершения хотя бы одной новой сделки.

Работа с сервисом MQL5 способствовала проведению экспериментов по формированию информационной среды мультитрейдинга, которые были начаты в феврале 2018 года. Статистические данные собирались на счетах брокера ForexClub, специализированных для стратегий скальпинга и свинг-трейдинга.

Видно, что к моменту завершения стадии эмпирического подбора индикаторов технического анализа и конфигурирования компонентов (12.08.2021) эффективность скальпинга и свинг-трейдинга достигла 77,8% и 68,9%, соответственно (см. Рис. 15).

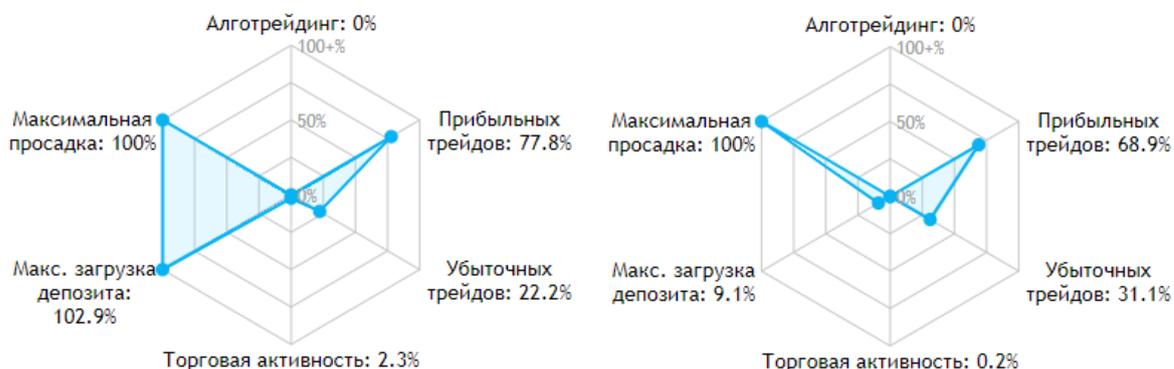


Рис. 15. Фрагменты экранных снимков веб-страниц с аналитическими данными торговых сигналов 935137 Intraday Scalper (слева) и 957125 Intraday Swinger (справа) для социального трейдинга на платформе MetaTrader 5 [14].

С. ДОПОЛНИТЕЛЬНЫЕ СЕРВИСЫ MQL5: ИНДИКАТОРЫ, СТАТЬИ, БИБЛИОТЕКА ИСХОДНЫХ КОДОВ, АЛГОТРЕЙДИНГ

Сообщество MQL5 предоставляет разнообразные возможности для совершенствования трейдинга [14]. На сервисе MQL5 находятся библиотеки исходных кодов терминалов и их компонентов (индикаторы, советники). В дополнение к 80 индикаторам технического анализа в комплекте поставки терминала MetaTrader 5 трейдер может найти на сервисе MQL5 множество новых индикаторов и пользовательских модификаций классических инструментов технического анализа. Советники (торговые роботы) предназначены для анализа ценовых графиков и автоматической торговли (алготрейдинга) на основе заложенных правил со сложными алгоритмами выработки сигналов и защитой от торговых ошибок. В зависимости от желания авторов индикаторы могут подключаться к терминалам трейдеров безвозмездно или посредством абонеента с оплатой через счёт сервиса MQL5.

D. АНАЛИТИЧЕСКАЯ ВИДЕОХРОНИКА

Флуктуации цены в процессе формирования одной свечи могут приводить

к существенным изменениям состояния индикаторов технического анализа, при которых создаются условия для появления ложных торговых сигналов. Для отбраковки ложных сигналов при разработке торговой системы мультитрейдинга полезно в реальном времени сопоставлять состояние всех индикаторов в окнах шести таймфреймов Аналитического дисплея. Ретроспективный анализ общей динамической картины совокупности индикаторов в торговом терминале осуществить невозможно, поскольку хронологическая база данных финансовых инструментов содержит сведения о состоянии индикаторов только на момент окончания формирования 1-минутных свечей. Таким образом, для разработки торговой системы мультитрейдинга и решения спорных вопросов со службой поддержки брокера при возникновении технических проблем в процессе трейдинга целесообразно осуществлять круглосуточную видеорегистрацию содержимого Аналитического и Обзорного дисплеев. Одновременная видеозапись тиковых графиков торговых платформ разных брокеров позволяет определить оптимальные моменты переключения торговых стратегий в соответствии со спецификой торговых сессий.

В результате экспериментов по оптимизации процесса длительной видеорегистрации на компьютерах с минимальной конфигурацией (оперативная память – 4 Gb, хранилище – 64 Gb) было установлено, что наивысшее соотношение надёжность/ресурсоёмкость при длительном полноэкранном захвате обеспечивают специализированные программы HyperCam, начиная с 4-й версии [28]. В опциях этой программы можно установить частоту кадров 1/сек с интервалом ключевых кадров 60–100 и компрессией в формат mp4 без звука (Encoding Algorithm: FFmpeg libx264, Preset: Very Slow, Profile: Baseline, Bitrate mode: Variable), что позволяет уменьшить до 1,5 Gb размер файла 24-часовой полноэкранной регистрации дисплея с размером FullHD (1920x1080). При видеозаписи содержимого Обзорного дисплея с разрешением 1200x1600 рекомендуется устанавливать интервал ключевых кадров на уровне 30 для предотвращения программных сбоев. При такой настройке размер файла круглосуточной видеозаписи Обзорного дисплея оказывается близок к 2,5 Gb. Таким образом, общий размер видеозаписей содержимого Аналитического и Обзорного дисплеев за 5 рабочих дней составляет 20 Gb, что в точности соответствует еженедельной квоте подписки Vimeo Pro (<https://vimeo.com/store/pro>).



Рис. 16. Экранный снимок кадра суточной видеозаписи содержимого настроенного для свинг-трейдинга Аналитического дисплея инструмента EURUSD с разметкой волн Эллиотта [20] на аккаунте видеохостинга Vimeo.com с сервисами подписки «Pro» в режиме работы с целевыми заметками. В ценовом окне таймфрейма H1 выделенной меткой обозначено начало 3-й волны Эллиотта. В Аналитических окнах применен шаблон 20210503.

Для обеспечения оперативного доступа к видеозаписям целесообразно экспортировать их на видеохостинги с учётом специфических особенностей. К примеру, видеохостинг YouTube.com не размещает видеозаписи продолжительностью 12 часов и более. В этом случае круглосуточные видеозаписи дисплеев мультитрейдинга целесообразно разделять надвое, используя видеоредактор HyperCam Video Editor, позволяющий сохранять файлы без перекодирования видеорядов. Начиная с сентября 2021 года, 12-часовые фрагменты видеозаписей Аналитического и Обзорного дисплеев мультитрейдинга (09:00-21:00) ежедневно публикуются в соответствующих плейлистах на видеоканале Multitrading – YouTube. Плейлист «Секреты» группирует видеозаписи компаративных экспериментов работы различных брокерских площадок, семинары, мастер-классы, видеоинструкции по работе с информационной средой мультитрейдинга.

Видеохостинг Vimeo.com не ограничивает продолжительность видеозаписей и предоставляет для подписчиков сервис целевых заметок [29], позволяющий отмечать и подписывать множество точек любого кадра видеоряда. На одном кадре может быть расположено более сотни меток. Адресами меток удобно делиться и использовать в веб-браузере для акцентирования внимания на сигналах технических индикаторов и волнового анализа (см. Рис. 16). Архив полных видеозаписей Аналитического дисплея размещён на хостинге Vimeo (<https://vimeo.com/showcase/analytical-display>). На сервисе Vimeo также находятся избранные видеозаписи Обзорного дисплея (<https://vimeo.com/showcase/overview-display>) и специальный канал <https://vimeo.com/channels/multitrading/>.

Е. ГРУППЫ СОЦИАЛЬНЫХ СЕТЕЙ

Для трансляции новостей системы мультитрейдинга, оперативной информации о ситуациях на финансовых рынках и обмена опытом в августе 2021 года были созданы группы в мессенджере Telegram (https://t.me/multitrading_pro) и в структуре популярных социальных сетей <https://vk.com/multitrading>, <https://www.facebook.com/multitrading.pro>.

Сервис Twitter (<https://twitter.com/multitrading4u>) оказалось невозможно использовать из-за отклонения любых публикаций: «Этот запрос похож на автоматизированные запросы от спам-бота. Чтобы защитить наших пользователей от спама и прочих вредоносных действий, мы отклоняем ваш запрос. Повторите попытку позже».

Сервис Инстаграм (<https://www.instagram.com/multitrading.pro/>) не обеспечивает автоматическую генерацию анонсовых изображений по данным гиперссылок в публикуемом сообщении. Оказалось, что удобно генерировать анонсовые изображения в формате GIF в сервисе Vimeo.com и подключать их к сообщениям Instagram. Эмпирическим путём была установлена оптимальная маршрутизация информации о медиаресурсах системы мультитрейдинга: при экспорте видеофайлов на видеохостинги Vimeo.com и YouTube.com артикул ресурса и релевантные ссылки регистрируются (см. 6.6) в менеджере оперативных задач Microsoft ToDo [30], прикрепляются к структуре ассоциативной карты TheBrain [31] и аннотируются там, после чего формируются сообщения в Telegram, которые ко-

пируются в VK.vom, систему Facebook.com+Instagram.com (через единый интерфейс Business Suite) и новостную ленту портала «Мультитрейдинг».

F. ПОРТАЛ «МУЛЬТИТРЕЙДИНГ»

Функцию основного сетевого ресурса осуществляет портал «Мультитрейдинг» (<https://multitrading.pro>), созданный на базе CMS 1С-Bitrix: Управление сайтом, редакция «Бизнес») в многодоменной системе информационного континуума [32]. Портал «Мультитрейдинг» содержит функционал собственной социальной сети (живая лента, бизнес-мессенджер, форум, блоги, опросы, база знаний), в которой основана группа «Мультитрейдер». К соответствующему разделу портала подключены видеоальбомы с записями дисплеев мультитрейдинга и мастер-классов. Раздел «Обучение» предназначен для размещения учебных курсов по тематике мультитрейдинга. Для тренировок и реальной практики трейдинга к структуре портала «Мультитрейдинг» подключён веб-терминал MetaTrader 5.

G. АДМИНИСТРИРОВАНИЕ ИНФОРЕСУРСОВ МУЛЬТИТРЕЙДИНГА

Оперативный учёт видеозаписей дисплеев мультитрейдинга, тактическое планирование действий и первичную фиксацию идей обеспечивает функционал менеджера задач Microsoft ToDo [30]. Аннотирование, систематизация и дополнение медиаресурсов релевантными связями осуществляются в ассоциативной карте TheBrain [31]. Для аккумуляции экранных снимков с примерами состояния технических индикаторов, анализа корреляций, формулировки закономерностей определения достоверных торговых сигналов, планирования экспериментов и анализа их результатов используется приложение Microsoft OneNote for Windows 10. Синхронизация шаблонов и релевантных данных на всех компьютерах и мобильных устройствах осуществляется через облачный сервис OneDrive.

7. ПРИНЦИПЫ РАЗРАБОТКИ ТОРГОВОЙ СИСТЕМЫ

Первые шаги на пути к профессиональному трейдингу – подбор и освоение торговой платформы, настройка оптимальной совокупности индикаторов технического анализа и разработка на основе их интерпретации собственной торговой системы (правил открытия, удержания и закрытия сделок) с учётом специфики финансовых инструментов и предпочтительной спекулятивной стратегии [25].

Интерпретация сигналов индикаторов и их использование в процессе разработки торговой системы облегчаются при сопоставлении фрагментов видеозаписей содержимого Аналитического дисплея со сходным изменением цен, которое предсказуемо происходит во время выхода новостей Nonfarm Payrolls (см. Рис. 17). В качестве примера рассмотрим возможную тактику оперативных действий в условиях высокоамплитудных изменений цены финансового инструмента EURUSD в минутном масштабе, которые наблюдались в 15:30 (MSK) 03.09.2021 и 08.10.2021 (см. Рис. 17).

На основании трёхлетних наблюдений было установлено, что высокоамплитудные ценовые скачки всегда происходят в направлении тренда старшего таймфрейма. Таким образом, индикаторы Аналитических окон старшего таймфрейма (M15 на Рис. 17) целесообразно использовать для определения направления ближайшего изменения цены. Видно, что 03.09.2021 к моменту ценового скачка трендовые (зеленые) линии Комплексных Индикаторов Тенденций в окнах упреждающих и подтверждающих индикаторов (метки 3 и 6 соответственно) направлены в сторону роста цены и пересекают уровень 50, что указывает на предстоящее ускорение изменения цены с высокой амплитудой. В Аналитическом окне таймфрейма M15 Комплексные Индикаторы Тенденций в окнах упреждающих и подтверждающих индикаторов (метки 9 и 12, соответственно) находятся выше уровня 50, но направлены в сторону снижения цены, что указывает на предстоящее увеличение цены с меньшей амплитудой, чем 03.09.2021.

Индикаторы Аналитических окон торгового таймфрейма (M4 на рис. 16) используются для прогнозирования амплитуды предстоящих осцилляций. В Аналитических окнах основных торговых таймфреймов заметно, что в момент скачка происходит разворот осцилляторов у направленных по предстоящему тренду линий индикаторов CITS (метки 2, 8 на Рис. 17) и CTR (метки 5, 11 на Рис. 17). На основе анализа видеохроник Аналитического дисплея статистически достоверно установлено, что подобная конфигурация индикаторов сигнализирует о начале 3 волны Эллиотта, которая по амплитуде превосходит все ближайшие осцилляции.



Рис. 17. Сопоставление кадров видеозаписей рабочих областей Аналитического дисплея с котировками инструмента EURUSD на момент выхода новостей NFP в 15:30(МСК) 03.09.2021 и 08.10.2021 (шаблон 20210812). Пронумерованные стрелки указывают на сигналы индикаторов технического анализа в Аналитических окнах с таймфреймами М1, М4, М15 (пересечение осциллятора Stoch и CITS: 1, 2, 3, 7, 8, 9; расположение и состояние индикатора CITS: 4, 5, 6, 10, 11, 12).

Индикаторы Аналитических окон с младшим таймфреймом (М1 на Рис. 17) применяются для определения оптимального момента открытия сделки. Видно, что ценовой скачок коррелирует с пересечением (метки 1 и 4 на Рис. 17) или разворотом (метки 7, 10 на Рис. 17) осцилляторов (Stoch, RSI) и соответствующих Комплексных Индикаторов Тенденций (CITS и CITSR) на уровне 50. Совпадение линий

подтверждающих индикаторов MFI (золотая линия) и CCI (голубая линия) в нижних частях Аналитических окон указывает на целесообразность открытия сделок с максимально доступным лотом в направлении, определённом по сигналам Упреждающих и Основных индикаторов.

Разработка торговой системы мультитрейдинга продолжается. Для создания универсальной торговой системы предстоит определить специфические сигналы выбора основного торгового таймфрейма, сформировать индивидуальные тактики работы в тренде и флэте, а затем – тактику смены тактик.

ЗАКЛЮЧЕНИЕ

Проект «Мультитрейдинг» стартовал 14 февраля 2018 года. Во время фазы инициализации, которая продолжалась до конца 2019 года, были определены цели, поставлены задачи, завершена оценка характеристик программно-аппаратной базы информационной среды мультитрейдинга и принято решение о целесообразности продолжения проекта. Эта статья подводит итоги фазы концептуализации системы мультитрейдинга (2020–2021 гг.), во время которой была сформулирована руководящая идея, установлен конструктивный принцип, определены перспективные задачи и намечены пути их решения, подготовлены информационные, материальные, финансовые и временные ресурсы. Следующая фаза – разработка и апробация торговой системы на базе руководящих принципов и информационной среды мультитрейдинга.

СПИСОК ЛИТЕРАТУРЫ

1. *Каспаринский Ф.О.* Информационная среда мультитрейдинга // Видеозапись доклада на XXIII Всероссийской научной конференции «Научный сервис в сети Интернет» 23 сентября 2021 г., онлайн). 2021. URL: <https://youtu.be/VUwxgOsKkhs>
2. *Савостина Л.С.* Брокер // Большая российская энциклопедия. Том 4. Москва, 2006. С. 223. URL: <https://bigenc.ru/economics/text/1884322>
3. *Дышлевский С.В.* Фундаментальный анализ // Большая российская энциклопедия. Том 33. Москва, 2017. С. 655–656. URL: <https://bigenc.ru/economics/text/4725550> .
4. *Дышлевский С.В.* Технический анализ // Большая российская

энциклопедия. Том 32. Москва, 2016. С. 110–111.

URL: <https://bigenc.ru/economics/text/4190812> .

5. Форекс Платформы 2021.

URL: <https://forex-ratings.ru/forex-trading-platforms/>

6. Форекс Брокеры. Рейтинг Брокеров Форекс 2021, основанный на отзывах реальных трейдеров рынка Forex с 2006 года. URL: <https://forex-ratings.ru>

7. Countries – Financial Action Task Force (FATF). 2021.

URL: <https://www.fatf-gafi.org/countries/>

8. Trading platform UTIP – forex platform to trade on the foreign exchange market // UTIP Technologies Ltd. 2021. URL: <https://www.utip.org>

9. Trading Desk Pro 5. Noveishaia torgovaia platforma dlia rynka Foreks dlia PC // Forex EuroClub. 2021. URL: <https://www.fxeuroclub.com/dd2000.php>

10. FxPro. Trade Like a Pro. FX брокер №1 в мире // FxPro Global Markets Limited 2021. URL: <https://www.fxproru.group>

11. Libertex Торговая платформа – программа для торговли на бирже // Forex Club International Limited. 2021. URL: <https://app.libertex.org>

12. Платформы для трейдинга и инвестиций // Forex Club International Limited. 2021. URL: <https://www.fxclub.org>

13. MetaQuotes – the developer of trading platforms for brokers, banks, exchanges and hedge funds // MetaQuotes Ltd. URL: <https://www.metaquotes.net>

14. MQL5: язык торговых стратегий для MetaTrader 5, позволяет писать собственные торговые роботы, технические индикаторы, скрипты и библиотеки функций // MQL5.community. URL: <https://www.mql5.com>

15. MetaTrader 5 – это единая удобная платформа для работы на рынке форекс // ООО Альфа-Форекс. 2021. URL: <https://alfaforex.ru/metatrader/>

16. Торговая платформа MetaTrader 5 – больше чем просто приложение для работы на FOREX // ООО ВТБ Форекс. 2021. URL: <https://vtbforex.ru/metatrader>

17. Торговая платформа MetaTrader 5 // Forex Club International Limited. 2021. URL: <https://go.libertex.com/visit/?bta=44551&nci=5360>

18. The Dow Theory | Schannep Timing Indicators // Schannep Timing Indicator & TheDowTheory.com Newsletter. URL: <https://thedowtheory.com>

19. Федеральный закон «О внесении изменений в Федеральный закон «О рынке ценных бумаг» и отдельные законодательные акты Российской

Федерации» от 31.07.2020 № 306-ФЗ (последняя редакция).

URL: http://www.consultant.ru/document/cons_doc_LAW_358796/

20. Expert Market Forecasting Using the Elliott Wave Principle // Elliott Wave International. 2021. URL: <https://www.elliottwave.com>

21. *Беляев Ю.И., Гербер Ю.В., Пророков А.Е., Котельников А.А., Беляева Е.Ю.* Осцилляторная модель прогноза флуктуации экономики // Вестник Международной академии системных исследований. Информатика, экология, экономика. 2015. Т. 17. № 1. С. 65–68.

URL: <https://www.elibrary.ru/item.asp?id=24898467>

22. *Каспаринский Ф.О.* Кинетика электрогенного транспорта двухвалентных катионов в митохондриях // Диссертация на соискание ученой степени кандидата биологических наук. Москва. МГУ имени М.В. Ломоносова. 2000. С. 118–120. URL: <https://istina.msu.ru/dissertations/20132285/>

23. *Каспаринский Ф.О., Полянская Е.И.* Интернет-активность учёного в рамках парадигмы инфоцентризма // Научный сервис в сети Интернет: труды XVII Всероссийской научной конференции (21–26 сентября 2015 г., г. Новороссийск). М.: ИПМ им.М.В. Келдыша, 2015. С. 141–149.

URL: <https://www.elibrary.ru/item.asp?id=24301153>

24. Допустимые величины психофизиологических производственных факторов по показателям тяжести и напряженности труда. Таблица 5.75 // Постановление Главного государственного санитарного врача Российской Федерации №2 от 28.01.2021 «Об утверждении санитарных правил и норм СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания». С. 368.

URL: https://www.rospotrebnadzor.ru/files/news/GN_sreda%20obitaniya_compressed.pdf

25. *Дышлевский С.В.* Спекулятивные стратегии // Большая российская энциклопедия. Том 31. Москва, 2016. С. 59.

URL: <https://bigenc.ru/economics/text/4246917>

26. *Elder A.* The New Trading For a Living // Education for intelligent traders. 2021. URL: <https://www.elder.com/product/elder-new-trading-for-a-living/>

27. Каспаринский Ф.О. Специализация использования микрокомпьютеров. // Электронные библиотеки. 2020. Т. 23, вып. 4. Тематический выпуск «Научный сервис в сети Интернет». Часть 2. С. 746–769.

URL: <https://elbib.ru/article/view/612/707>

28. Hypercam – video zakhvat ekrana // Solveig Multimedia, 2021.

URL: <https://www.solveigmm.com/ru/products/hypercam/>

29. Каспаринский Ф.О., Полянская Е.И. Информационно-навигационный сервис сетевых аудиовизуальных ресурсов // Научный сервис в сети Интернет: труды XX Всероссийской научной конференции (17–22 сентября 2018 г., г. Новороссийск). М.: ИПМ им. М.В. Келдыша, 2018. С. 284–294.

<https://doi.org/10.20948/abrau-2018-42>

30. Каспаринский Ф.О. Тактическая сортировка управленческих задач при их администрировании посредством меток Приоритетов, Спецификаций и Аффилиаций // Электронные библиотеки. 2020. Т. 23, вып. 4. Тематический выпуск «Научный сервис в сети Интернет». Часть 2. С. 733–745.

<https://doi.org/10.26907/1562-5419-2020-23-4-733-745> .

31. Каспаринский Ф.О. Администрирование информационной среды посредством ассоциативной карты TheBrain 9 // Научный сервис в сети Интернет: труды XX Всероссийской научной конференции (17–22 сентября 2018 г., г. Новороссийск). М.: ИПМ им. М.В. Келдыша, 2018. С. 275–283.

<https://doi.org/10.20948/abrau-2018-40>

32. Каспаринский Ф.О. Инфоконтинуум как сервис для междисциплинарной системной интеграции профессиональных интернет-представительств // Научный сервис в сети Интернет: труды XVIII Всероссийской научной конференции (19–24 сентября 2016 г., г. Новороссийск). М.: ИПМ им. М.В. Келдыша, 2016. С. 162–169.

<https://doi.org/10.20948/abrau-2016-12>

PRINCIPLES OF MULTITRADING

F. O. Kasparinsky^[0000-0002-1048-9212]

MASTER-MULTIMEDIA LLC, Entuziastov Shosse 98-3-274, Moscow, 111531

felix@kasparinsky.pro

Abstract

Modern software and hardware tools provide unprecedented freedom for a variety of activities in the forex markets, from trading to analyzing the feasibility of models of nonlinear processes in self-organizing systems. To reduce risks and increase the efficiency of interaction with stock market instruments, it is proposed to provide variable adaptability of trading by combining trading strategies using several trading accounts of different brokers, multiple financial instruments, and Complex Indicators Tendencies of price changes. As a result of three years of experimental work, the basic principles of multitrading have been formulated and tested, and an information environment has been compiled, contributing to the development of an individualized trading system. The basic concept of organizing a multitrading information environment: the use of specialized hardware and software systems for strategic analysis and forecasting of price changes for an individual financial instrument, tactical selection of a promising financial instrument from the available set, and effective operating activities with orders of trading accounts. It can be expected that the evolution of the principles of multitrading will lead to the creation of analytical systems for predicting the kinetics of non-equilibrium changes in the characteristic parameters of self-organizing cooperative systems for wide application in biology, cybernetics, economics, and the social sphere.

Keywords: *multitrading, trading, forex, technical analysis, investments, work organization, efficiency, financial market, oscillations, forecast.*

REFERENCES

1. Kasparinsky F.O. Informatsionnaia sreda multitreidinga // Videozapis doklada na XXIII Vserossiiskoi nauchnoi konferentsii «Nauchnyi servis v seti Internet» 23 sentiabria 2021 g., onlain). 2021. URL: <https://youtu.be/VUwxgOsKkhs>

2. Savostina L.S. Broker // Bolshaia rossiiskaia entsiklopediia. Tom 4. Moskva,

2006. S. 223. URL: <https://bigenc.ru/economics/text/1884322>

3. *Dyshlevskii S.V.* Fundamentalnyi analiz // Bolshaia rossiiskaia entsiklopediia. Tom 33. Moskva, 2017. S. 655–656. URL: <https://bigenc.ru/economics/text/4725550>.

4. *Dyshlevskii S.V.* Tekhnicheskii analiz // Bolshaia rossiiskaia entsiklopediia. Tom 32. Moskva, 2016. S. 110–111. URL: <https://bigenc.ru/economics/text/4190812>

5. Foreks Platformy 2021. URL: <https://forex-ratings.ru/forex-trading-platforms/>

6. Foreks Brokery. Reiting Brokerov Foreks 2021, osnovannyi na otzyvakh realnykh treiderov rynka Forex s 2006 goda. URL: <https://forex-ratings.ru>

7. Countries – Financial Action Task Force (FATF). 2021. URL: <https://www.fatf-gafi.org/countries/>

8. Trading platform UTIP – forex platform to trade on the foreign exchange market // UTIP Technologies Ltd. 2021. URL: <https://www.utip.org>

9. Trading Desk Pro 5. Noveishaia torgovaia platforma dlia rynka Foreks dlia PC // Forex EuroClub. 2021. URL: <https://www.fxeuroclub.com/dd2000.php>

10. FxPro. Trade Like a Pro. FX broker №1 v mire // FxPro Global Markets Limited 2021. URL: <https://www.fxproru.group>

11. Libertex Torgovaia platforma – programma dlia trgovli na birzhe // Forex Club International Limited. 2021. URL: <https://app.libertex.org>

12. Platformy dlia treidinga i investitsii // Forex Club International Limited. 2021. URL: <https://www.fxclub.org>

13. MetaQuotes – the developer of trading platforms for brokers, banks, exchanges and hedge funds // MetaQuotes Ltd. URL: <https://www.metaquotes.net>

14. MQL5: language of trade strategies built-in the MetaTrader 5 Trading Platform, allows writing your own trading robots, technical indicators, scripts and libraries of functions. URL: <https://www.mql5.com>

15. MetaTrader 5 – eto edinaia udobnaia platforma dlia raboty na rynke foreks // OOO Alfa-Foreks. 2021. URL: <https://alfaforex.ru/metatrader/>

16. Torgovaia platforma MetaTrader 5 – bolshe chem prosto prilozhenie dlia raboty na FOREX // OOO VTB Foreks. 2021. URL: <https://vtbforex.ru/metatrader>

17. Torgovaia platforma MetaTrader 5 // Forex Club International Limited. 2021. URL: <https://go.libertex.com/visit/?bta=44551&nci=5360>

18. The Dow Theory | Schannep Timing Indicators // Schannep Timing Indicator & TheDowTheory.com Newsletter. URL: <https://thedowtheory.com>

19. Federalnyi zakon "O vnesenii izmenenii v Federalnyi zakon "O rynke tsennykh bumag" i otdelnye zakonodatelnye akty Rossiiskoi Federatsii" ot 31.07.2020 N 306-FZ. URL: http://www.consultant.ru/document/cons_doc_LAW_358796/

20. Expert Market Forecasting Using the Elliott Wave Principle // Elliott Wave International. 2021. URL: <https://www.elliottwave.com>

21. *Beliaev Iu.I., Gerber Iu.V., Prorokov A.E., Kotelnikov A.A., Beliaeva E.Iu.* Ostsillatornaia model prognoza fluktuatsii ekonomiki // Vestnik Mezhdunarodnoi akademii sistemnykh issledovani. Informatika, ekologiya, ekonomika. 2015. T. 17. № 1. S. 65–68. URL: <https://www.elibrary.ru/item.asp?id=24898467>

22. *Kasparinsky F.O.* Kinetika elektrogenogo transporta dvukhvalentnykh kationov v mitokhondriakh // Dissertatsiia na soiskanie uchenoi stepeni kandidata biologicheskikh nauk. Moskva. MGU imeni M.V. Lomonosova. 2000. S. 118–120. URL: <https://istina.msu.ru/dissertations/20132285/>

23. *Kasparinsky F.O., Polyanskaya E.I.* Internet-aktivnost uchenogo v ramkakh paradigmy infotsentrizma // Nauchnyi servis v seti Internet: trudy XVII Vserossiiskoi nauchnoi konferentsii (21–26 sentiabria 2015 g., g. Novorossiisk). M.: IPM im. M.V. Keldysha, 2015. S. 141–149. URL: <https://www.elibrary.ru/item.asp?id=24301153>

24. Dopustimye velichiny psikhofiziologicheskikh proizvodstvennykh faktorov po pokazateliam tiazhesti i napriazhennosti truda. Tablitsa 5.75 // Postanovlenie Glavnogo gosudarstvennogo sanitarnogo vracha Rossiiskoi Federatsii №2 ot 28.01.2021 «Ob utverzhdenii sanitarnykh pravil i norm SanPiN 1.2.3685-21 «Gigienicheskie normativy i trebovaniia k obespecheniiu bezopasnosti i (ili) bezvrednosti dlia cheloveka faktorov sredy obitaniia». S.368. URL: https://www.rospotrebnadzor.ru/files/news/GN_sreda%20_obitaniya_compressed.pdf

25. *Dyshlevskii S.V.* Spekuliativnye strategii // Bolshaia rossiiskaia entsiklopediia. Tom 31. Moskva, 2016. S. 59. URL: <https://bigenc.ru/economics/text/4246917>

26. *Elder A.* The New Trading For a Living // Education for intelligent traders. 2021. URL: <https://www.elder.com/product/elder-new-trading-for-a-living/>

27. *Kasparinsky F.* Configuring the Information Environment of Microcomputers with the Microsoft Windows 10 Operating System. In: CEUR Workshop Proceedings,

vol. 2543, pp. 185–198. 2020. URL: <http://ceur-ws.org/Vol-2543/rpaper17.pdf>

28. Hypercam – video zakhvat ekrana // Solveig Multimedia, 2021. URL: <https://www.solveigmm.com/ru/products/hypercam/>

29. *Kasparinsky F.O., Polyanskaya E.I.* Informatsionno-navigatsionnyi servis setevykh audiovizualnykh resursov // Nauchnyi servis v seti Internet: trudy XX Vserossiiskoi nauchnoi konferentsii (17–22 sentiabria 2018 g., g. Novorossiisk). M.: IPM im. M.V. Keldysha, 2018. S. 284–294. <https://doi.org/10.20948/abrau-2018-42> .

30. *Kasparinsky F.* Using the Labels of Priorities, Specifications, and Affiliations when Working in Task Management Programs. In: CEUR Workshop Proceedings, vol. 2543, pp. 399–406. 2020. URL: <http://ceur-ws.org/Vol-2543/spaper11.pdf>

31. *Kasparinsky F.O.* Administrirovanie informatsionnoi sredy posredstvom assotsiativnoi karty TheBrain 9 // Nauchnyi servis v seti Internet: trudy XX Vserossiiskoi nauchnoi konferentsii (17-22 sentiabria 2018 g., g. Novorossiisk). M.: IPM im. M.V. Keldysha, 2018. S. 275–283. <https://doi.org/10.20948/abrau-2018-40>

32. *Kasparinsky F.O.* Infokontinuum kak servis dlia mezhdistsiplinarnoi sistemnoi integratsii professionalnykh internet-predstavitelstv // Nauchnyi servis v seti Internet: trudy XVIII Vserossiiskoi nauchnoi konferentsii (19–24 sentiabria 2016 g., g. Novorossiisk). M.: IPM im. M.V. Keldysha, 2016. S. 162–169. <https://doi.org/10.20948/abrau-2016-12> .

СВЕДЕНИЯ ОБ АВТОРЕ



КАСПАРИНСКИЙ Феликс Освальдович – кандидат биологических наук, основатель и научный руководитель Лаборатории мультимедийных технологий Биологического факультета МГУ имени М.В. Ломоносова, учредитель и Генеральный директор ООО «МАСТЕР-МУЛЬТИМЕДИА», представляющий брокер (Introducing Broker, IB) международной компании Forex Club. Сфера научных интересов: оптимизация информационной среды, создание персонального информационного континуума.

Felix Oswaldovich KASPARINSKY – Candidate of Biological Sciences, Founder and Scientific Director of Multimedia Technologies Laboratory (Biological Faculty, M.V. Lomonosov Moscow State University), Founder and General Director of MASTER-MULTIMEDIA LLC., Introducing Broker (IB) of the Forex Club International LLC. Research interests: optimization of the information environment, creation of a personal information continuum.

email: felix@kasparinsky.pro;

ORCID: 0000-0002-1048-9212

Материал поступил в редакцию 25 октября 2021 года

УДК 004.021, 004.42

МЕТОД ПОИСКА ЭКСПЕРТОВ ПО ДАННЫМ НАУКОМЕТРИЧЕСКИХ СИСТЕМ

А. С. Козицын¹ [0000-0002-8065-9061], С. А. Афонин² [0000-0003-3058-9269]

^{1,2} НИИ механики МГУ им. М.В. Ломоносова, Мичуринский пр., 1, Москва, 119192

¹alexanderkz@mail.ru, ²serg@msu.ru

Аннотация

Применение современных методов тематического анализа для аналитической обработки больших объемов информации используется в настоящее время практически во всех сферах человеческой деятельности, в том числе, в наукометрии. Многие наукометрические системы и системы цитирования, включая всемирно известные WoS, Scopus, Google Scholar, разрабатывают тематические рубрикаторы для поиска и обработки информации. Важными практическими задачами, которые могут решаться с применением методов тематической классификации, являются: оценка динамики развития тематических направлений в организации, отдельной стране и мировой науке в целом; поиск статей по заданной тематике; поиск и оценка авторитетности экспертов; поиск журналов для публикации и другие актуальные задачи. Авторами созданы программные реализации алгоритмов для решения некоторых из перечисленных задач и ведутся научные исследования с целью создания новых эффективных математических моделей и алгоритмов в этой области.

Ключевые слова: тематический поиск, библиографические данные, поиск экспертов, информационные системы, наукометрия.

ВВЕДЕНИЕ

С развитием средств коммуникаций и систем удаленного взаимодействия все большую актуальность приобретает задача поиска высококвалифицированных специалистов и экспертов для выполнения работ в удаленном режиме. Такие работы могут являться краткосрочными, периодическими или выполняться на постоянной основе. В настоящее время поиск экспертов наиболее востребован в СМИ. Журналисты и издатели производят поиск консультантов по тематикам пуб-

лируемых статей. Владельцы новостных и информационных порталов ищут специалистов для написания статей по заданным тематическим направлениям. В производственной сфере и при разработке программного обеспечения также производится поиск экспертов для проведения разовых консультаций, для работы на постоянной основе в стартапах и развивающихся проектах. Подобный поиск производится обычно через социальные сети или биржи задач. Активный поиск экспертов через социальные сети производится при помощи публикации объявлений и поиска среди знакомых или в специализированных сообществах, например, сообществ [1, 2], содержащих по 12 тысяч участников. Для СМИ также существуют биржи статей, на которых можно найти исполнителя для написания статьи по заданной тематике [3–5]. Поиск специалистов на таких биржах производится путем размещения заказов. Заказчики размещают задания, а эксперты выкладывают описания своих компетенций. Специалисты должны быть зарегистрированы на бирже и регулярно просматривать заказы по своей тематике. Такой метод поиска эффективен для поиска специалистов не очень высокой квалификации, которые специализируются на написании текстов для СМИ и интернет-ресурсов с целью заработка. Квалификация специалистов оценивается в большинстве случаев по количеству проданных заказов. Для успешного проведения поиска в таких системах необходимо активное участие в процессе поиска как заказчиков работ, так и самих экспертов.

Существуют специализированные сайты, которые позволяют экспертам обмениваться контактами. Примером служит проект deadline.media, предназначенный для поиска специалистов для СМИ. Каждый зарегистрированный пользователь самостоятельно отмечает темы, в которых он является экспертом, и его контакты будут высвечиваться при поиске по рубрикам. Рубрикатор одноуровневый, содержит около 80 рубрик, например, «наука, инновации и технологии», «происшествия», «культура и искусство». База контактов содержит около 14 тысяч пользователей разной степени актуальности. Также можно упомянуть следующие англоязычные сервисы: www.profnetconnect.com для поиска PR-специалистов; свернутый на настоящий момент проект поиска экспертов в произвольных областях www.allexperts.com; сервис для анализа материалов экспертами-добровольцами

перед публикацией pressfeed.ru. Заинтересованность экспертов в подобной работе может быть не только финансовая. Например, эксперт может быть заинтересован в дополнительном упоминании своего имени в СМИ.

Аналогичные биржи создаются для поиска эксперта при создании проектов, например, сайт для поиска экспертов expertme.ru. Организатор проекта указывает описание проекта и требования к эксперту, а эксперты при поиске указывают свои компетенции и просматривают предложения заказчиков. Оценка квалификации эксперта производится только на основании предоставленного им описания и последующего личного собеседования.

Общими недостатком представленных выше ручных и автоматизированных методов являются необходимость активного участия в поиске обеих сторон и отсутствие критериев оценки квалификации эксперта. Оценка производится либо на основе субъективной оценки небольшого количества других людей, либо на основе активности выполнения работ на поисковом сайте, что никак не связано с реальной квалификацией эксперта.

Задача тематического поиска экспертов в научной сфере существенно отличается от описанных выше задач поиска для СМИ и разработки ПО. Такой поиск востребован при подборе рецензентов для научных статей, выборе экспертов при проведении конкурсов, поиске высококвалифицированных консультантов для научной работы и выборе студентами научных руководителей. В большинстве случаев при подобном поиске эксперты не готовы принимать активное участие в процессе поиска. Ученый с мировым именем может согласиться прочитать и дать рецензию на статью по просьбе редакции, но он не будет самостоятельно искать направленные в редакцию для публикации статьи, на которые он мог бы дать рецензии. В этой связи применение биржевых систем для тематического поиска в научной сфере практически невозможно, и требуется использовать иные принципы и механизмы поиска.

1. ОБЗОР СУЩЕСТВУЮЩИХ СИСТЕМ

Наиболее простым способом поиска специалистов является поиск по описаниям компетенций в профиле сотрудника. Например, в проекте поиска экспертов среди сотрудников университетов США и Канады [6] осуществляется поиск по

ключевым словам среди сотрудников университетов США и Канады, зарегистрированных в системе. Контекстный поиск производится только по описаниям компетенций, которые представили сами сотрудники в описании своего профиля. Какой-либо оценки соответствия заявленной тематики и реальной научной деятельности не производится. Экспертам предлагается самостоятельно в личном профиле заполнить список своих компетенций, по которому в дальнейшем поисковая система будет осуществлять поиск по ключевым словам. Такой подход прост в реализации, однако имеет очень низкую полноту поиска. В своих компетенциях сотрудники указывают общие понятия, но не указывают специальных терминов. Например, в системе [6] можно найти 40 зарегистрированных экспертов с компетенцией “big data”, но нельзя найти ни одного специалиста по “k-means” или “SVM”. Кроме того, подобный метод не позволяет оценить квалификацию экспертов и провести ранжирование результатов поиска.

Существует значительное количество социальных сетей исследователей, позволяющих загружать свои публикации, обмениваться ими и осуществлять поиск. Проект Academia.edu позволяет загружать в профиль свои статьи и производить поиск по чужим статьям (полноценный полнотекстовый поиск доступен только на платной основе). Проект ResearchGate дополнительно организует совместную работу над статьями. Проект www.bibsonomy.org позволяет загружать и индексировать русскоязычные статьи, однако, в очень небольшом объеме. За 10 лет в нем содержится только 14 русскоязычных статей, содержащих слово «поиск».

Более качественные методы поиска экспертов реализуются на основе анализа данных о публикациях. Такие данные содержатся в системах цитирования и наукометрических системах. Проекты, реализующие подобный метод поиска, можно разделить на три группы.

Проекты первой группы позволяют производить тематический поиск публикаций с последующим ручным анализом информации об авторах. Например, проект CiteSeerX [7] университета штата Пенсильвания содержит около 1 млн документов и позволяет искать статьи по ключевым словам и фамилиям авторов. В профиле каждого автора предоставляется подборка его статей. Проект Seman-

ticscholar [8] содержит 40 млн статей и позволяет осуществлять поиск по ключевым словам. Сайты издательств предоставляют возможность поиска по ключевым словам и библиографическим данным для платного доступа к текстам статей [9–11]. Недостатком таких проектов является отсутствие оценки релевантности автора поисковому запросу. Пользователю необходимо найти статьи, соответствующие запросу, и вручную перебрать всех соавторов для оценки их компетенции.

Проекты второй группы не только позволяют проводить поиск по статьям, но и показывают агрегированные данные по авторам публикаций. Проект Digital Library [12] позволяет производить поиск статей по ключевым словам и показывает при просмотре результатов список авторов всех найденных статей с указанием для каждого автора количества статей. Имеет удобный интерфейс поиска, однако база поиска составляет всего 3 млн статей.

Проект Google Scholar [13] имеет наиболее полный индекс по статьям, в том числе на русском языке. При поиске по ключевым словам показывает количество цитирований статей по своим данным, значительно превышающее количество цитирований по другим системам. Следуя глобальной концепции Google по обучению системы силами самих пользователей, сервис предоставляет пользователю возможность создавать свой профиль и добавлять в профиль статьи из результатов выполнения поискового запроса или с заполнением формы описания статьи вручную. Таким образом, пользователь может добавить в свой профиль и русскоязычные статьи, но информация не верифицируется, и для большинства русскоязычных авторов профили отсутствуют или не заполнены. Оценка соответствия автора тематике запроса также производится вручную перебором всех соавторов найденных работ.

Проект Microsoft Academic Graph (academic.microsoft.com) [14] позволяет проводить поиск экспертов на основе анализа научных публикаций. На основе ключевых слов поискового запроса пользователю предоставляется подборка статей, релевантных запросу. Для сформированной подборки пользователю показываются распределения: по списку наиболее популярных тем для этих статей; по типам публикаций; журналам; научным организациям; конференциям; разделам классификаторов и авторам. Каждый элемент этих списков содержит показатель агрегированной релевантности запросу пользователя и может быть использован

для дополнительной фильтрации. Таким образом, можно выбрать автора или организацию, которые наиболее близки запросу пользователя, и просмотреть список их публикаций. Проект англоязычный, не обрабатывает запросы на русском языке и не анализирует русскоязычные статьи. Коррекция авторами данных о своих работах не предусмотрена.

К третьей группе относятся проекты с возможностью поиска профилей авторов по ключевым словам. Наиболее известным проектом этой группы является китайский проект AMiner [15]. Для проведения тематического анализа в нем используются данные о публикациях авторов, полученных патентах и участии в проектах. Поискový запрос на основе выбранных ключевых слов может быть уточнен с использованием одной из сорока рубрик дополнительного тематического классификатора. Список найденных авторов может ранжироваться как по соответствию тексту запроса, так и по различным показателям авторитетности автора. Для каждого автора в списке результатов поиска дополнительно показывается общее количество зарегистрированных статей автора, количество найденных цитирований и перечень автоматически построенных компетенций. Для детализации и дополнительного ознакомления с компетенциями автора есть возможность перейти в его автоматически построенный профиль, в котором представлены: распределение выделенных компетенций по годам публикаций; список соавторов; список публикаций, патентов и проектов автора. В отдельных вкладках для просмотра доступен перечень статей, проектов и патентов эксперта. Представлены соседние вершины из графа соавторства (Ego Network). Дополнительно представлено распределение количества статей с заданными ключевыми словам по годам. Например, можно видеть, что популярность “K-means” и “Elliptic curves” в открытых научных статьях была довольно высока в период с 2001 по 2011 годы, а потом начала спадать. А популярность “SVM”, наоборот, начала расти после 2008 года.

Наиболее серьезным недостатком перечисленных выше проектов с точки зрения русскоязычных пользователей является полное отсутствие или крайне незначительный объем русскоязычных статей, используемых для анализа. Этот факт особенно важен для экспертов в гуманитарных областях, поскольку результаты научной деятельности многих известных российских ученых гуманитарного

направления мало представлены в англоязычных журналах, индексируемых Web of Science.

Кроме того, проблемы с поддержкой русского языка влекут за собой сложности с определением имен и фамилий в англоязычных статьях, что приводит к некорректному созданию профилей и неправильной привязке статей к авторам. Например, для статьи [16] системой определены авторы “V. Valery, Z. Maxim, K. Alexander, K. Maxim, S. Dmitry” [17]. Также некорректно работает во многих случаях определение организаций. Например, для статьи [18] из пяти соавторов, работающих в Московском государственном университете им. М.В. Ломоносова, статьи трех авторов отнесены к организации “Moscow State University”, а двух – к “Michigan State University” [19]. Подобные ошибки определения авторов приводят к значительному ухудшению качества поиска среди русскоязычных авторов. Еще одним недостатком является отсутствие удобной фильтрации по территориальному признаку или языку. Во многих случаях пользователю требуется найти не лучшего эксперта в мире по выбранной тематике, а достаточно хорошего эксперта, с которым он сможет проконсультироваться.

Следует отметить, что проекты AMiner и Microsoft Academic Graph предоставляют пользователям возможность скачать свои данные для самостоятельного анализа: библиографические описания статей (авторы, название, журнал, страницы и другие); ключевые слова; ссылки на pdf-версию и аннотации (при наличии), а также описания авторов. Данные предоставляются только за 2018-й год. Данные последних лет, наиболее востребованные для проведения анализа, являются закрытыми.

Отдельно следует выделить русскоязычный проект РИНЦ [20]. В проекте не предусмотрен поиск авторов по ключевым словам, однако возможен поиск авторов по фамилии, различным идентификаторам (ORCID и другие), городу или организации. Дополнительно предусмотрена фильтрация по одной из 70 тематических рубрик. Результаты поиска ранжируются по цитируемости. Основным недостатком поискового модуля в этом проекте является невозможность конкретизации тематической области поиска, поскольку имеющиеся рубрики дают слишком общую формулировку темы (например, «математика», «геология»). Кроме того, в системе отсутствует возможность проведения оценки релевантности поисковому

запросу. При поиске по разделу «математика» эксперт с количеством цитирований 30 будет иметь более низкий ранг, чем эксперт с количеством цитирований 40, даже если у первого из них все работы посвящены математике, а у второго только 10% относятся к разделу «математика», а остальные – к другим разделам классификатора.

На основе представленного выше обзора существующих систем и оценки их сильных и слабых сторон можно сформулировать требования, которые предъявляются к системам тематического поиска экспертов в научной сфере. Система должна производить поиск в автоматическом режиме без участия в процессе поиска самих экспертов. Для обеспечения полноты поиска и более объективной оценки квалификации автора должны использоваться не только статьи, но и другая информация о научной деятельности эксперта. Предназначенная для использования в России система должна уметь работать с русскоязычными материалами, в том числе распознавать русскоязычные имена и фамилии. В системе должна обеспечиваться возможность выполнения поисковых запросов по общим тематическим разделам рубрикатора и специальным тематическим направлениям, например, с использованием для поиска заданных пользователем ключевых слов [21]. Результаты поиска должны автоматически ранжироваться по авторитетности найденных специалистов в заданной тематической области. Исходные данные для осуществления такого поиска можно получать из наукометрических систем.

2. ОПИСАНИЕ МЕТОДА

Использование информации из наукометрических систем [22, 23] позволяет производить автоматический поиск экспертов и ранжировать результаты поиска. Для тематического поиска авторов могут использоваться следующие данные, размещенные в наукометрической системе: библиографические описания статей, тезисов и монографий; доклады на конференциях; диссертации; описания проектов; патенты и свидетельства о регистрации ПО; выступления в СМИ; описание читаемых учебных курсов и другие. По каждому из перечисленных типов научной продукции в наукометрической системе может содержаться полнотекстовое описание (название, аннотация, ключевые слова) и список авторов, задаваемый вручную или с использованием автоматических алгоритмов распознавания [24, 25].

Для некоторых типов работ в описании может содержаться дополнительная информация, отражающая значимость работ: цитируемость, принадлежность спискам WoS и ВАК, количество сделанных по проекту публикаций, тип читаемых курсов, объем финансирования проектов и другие.

Задача тематического поиска сотрудников может быть сформулирована следующим образом. Заданы множество возможных запросов Q , множество авторов A и множество объектов, разбитых на непересекающиеся подмножества $S = S_1 \cup S_2 \cup \dots \cup S_N$, где N – количество типов научной продукции. Заданы функция релевантности текстового описания запросу поиска по ключевым словам и рубрикам $F: S \times Q \rightarrow \mathfrak{R}$, функция значимости работы $L: S \rightarrow \mathfrak{R}$ и функция принадлежности работы автору $P: S \times A \rightarrow (0..1)$. Требуется построить функцию соответствия автора поисковому запросу $L: A \times Q \rightarrow \mathfrak{R}$.

Функция F реализуется с использованием стандартных средств контекстного поиска с учетом морфологического анализа. Для улучшения полноты поиска может также использоваться автоматическое расширение набора ключевых слов близкими терминами и переводами [26].

Реализация функция L различна для разных типов работ. Для статей ее значение определяется вхождением журнала в списки ВАК, Scopus, РИНЦ, рейтингом журнала и цитируемостью самой статьи. Для проектов значение функции определяется объемом финансирования, а также количеством и значимостью сделанных по проекту публикаций. Для читаемых курсов учитывается категория курса: потоковые лекции; семинары; спецкурсы по выбору кафедры; спецкурсы по выбору студента. Для остальных типов работ значение функции тождественно равно 1.

Функция P отражает вероятную степень участия автора в создании работы. Как показано в работе [27], при определении функции дискретным образом $P: S \times A \rightarrow \{0,1\}$ (1 – для всех соавторов работы и 0 в других случаях) значительно падает точность поиска, поскольку соавторы могут только косвенно участвовать в написании работы и не являться специалистами по тематике, рассматриваемой в работе. Поэтому для задания связи авторов с публикациями и докладами используется формула из работы [25]. Для первого автора публикации значение функции

равно $\frac{1}{2} + \frac{1}{2n}$, для остальных авторов $\frac{1}{2n}$, где n – количество соавторов публикации. Для остальных типов объектов функция равна $\frac{1}{n}$.

На первом этапе выполнения запроса осуществляется поиск по объектам всех типов с определением релевантности запроса в соответствии с функцией релевантности текстового описания запросу. На основе полученной выборки для каждого типа объектов по каждому автору вычисляется коэффициент соответствия запросу

$$R_i(a, q) = \sum_{s \in S_i} F(s, q)P(s, a)L(s).$$

С использованием рассчитанных коэффициентов производится ранжирование авторов для каждого типа работ, и для каждого автора определяется его ранг $ord_i(a, q) \in (0..1)$. Значение 1 соответствует максимальному значению R_i по всем авторам для данного типа работ, 0 – минимальному. Суммарный рейтинг каждого автора вычисляется по формуле

$$M(a, q) = \sum k_i ord_i(a, q),$$

где k_i – коэффициенты значимости каждого типа, рассчитанные на этапе обучения.

Для всех отобранных авторов строится граф соавторства, в котором определяется авторитетность вершин S с использованием алгоритма PageRank [27]. Функция L определяется по формуле

$$L = M\sqrt{S}.$$

Разработанный алгоритм может использовать данные любой наукометрической системы. Однако при разработке программной реализации алгоритма необходимо учитывать формат данных, используемых в конкретной наукометрической системе. Разработанная авторами программная реализация алгоритма апробировалась на данных наукометрической информационно-аналитической системы ИСТИНА [22], которая содержит данные по 27 организациям, в том числе, МГУ им. М.В. Ломоносова, РХТУ им. Д.И. Менделеева, ИНЭОС РАН, ИФХЭ РАН, ФГАУ «НМИЦ Нейрохирургии им. ак. Н.Н. Бурденко» Минздрава России, МПГУ, ИФЗ РАН, ИНХС РАН, ИГ РАН и другие.

3. ИНТЕРФЕЙС ПОИСКА

Для просмотра результатов поиска и апробации алгоритма реализован интерфейс, позволяющий производить поиск по ключевым словам с возможностью фильтрации результатов по одной из 70 общих тематических рубрик. На рис. 1 приведен результат выполнения поискового запроса по ключевым словам «тематический анализ»

The screenshot shows a search interface with the following elements:

- Search filters:**
 - Ключевые слова (через запятую):
 - Область знаний:
 - Подразделение:
- Search results table:**

V	NAME	DEP
1.139	д. ф.-м. н. проф. Васенин Валерий Александрович	404 Лаборатория автоматизации экспериментальных исследований(Совм)(МГУ); Межфакультетская кафедра математического моделирования и компьютерных исследований(МГУ); Кафедра вычислительной математики(Совм)(МГУ) Заведующий лабораторией, Заведующий кафедрой, Профессор
1.012	к. ф.-м. н. Афонин Сергей Александрович	404 Лаборатория автоматизации экспериментальных исследований(МГУ); Кафедра вычислительной математики(Совм)(МГУ); Межфакультетская кафедра математического моделирования и компьютерных исследований(Совм)(МГУ) Ведущий научный сотрудник, Доцент, Доцент
0.957	к. ф.-м. н. Козицын Александр Сергеевич	404 Лаборатория автоматизации экспериментальных исследований(МГУ); Кафедра вычислительной математики(Совм)(МГУ) Ведущий научный сотрудник, Доцент
0.542	Щачнев Дмитрий Алексеевич	404 Лаборатория автоматизации экспериментальных исследований(Совм)(МГУ) Специалист (зоолог, программист, геолог, инженер и т.п.)
0.417	к. ф.-м. н. Голомазов Денис Дмитриевич	None
0.391	д. ф.-м. н. проф. Мальковский Михаил Георгиевич	Кафедра алгоритмических языков(МГУ) Профессор
0.314	к. т. н. Коршунов Андрей Александрович	404 Лаборатория автоматизации экспериментальных исследований(МГУ) Ведущий научный сотрудник
0.307	Занчурич Максим Анатольевич	404 Лаборатория автоматизации экспериментальных исследований(МГУ); Межфакультетская кафедра математического моделирования и компьютерных исследований(Совм)(МГУ) Младший научный сотрудник, Ассистент

Рис. 1. Результат поиска по запросу «Тематический анализ»

На рис. 2 приведен пример выполнения запроса по ключевым словам «Теория упругости».

Для детального изучения результатов поисковой выдачи можно перейти по ссылке на профиль автора или на список найденных по запросу научных результатов (рис. 3).

Разработанный интерфейс позволяет получить ранжированный список авторов, имеющих научные результаты в тематической области, заданной запросом, и провести более детальный анализ, в том числе ознакомится с аннотациями и полными текстами работ.

Ключевые слова (через запятую)	теория упругости	
Область знаний	Все	
Подразделение	Все	

1.167	д.ф.-м.н. проф. Шешенин Сергей Владимирович	Кафедра теории пластичности(МГУ) Профессор
1.028	д.т.н. Лурье Сергей Альбертович	Лаборатория механики прочности и разрушения материалов и конструкций(Совм.)(ИПМех РАН); Кафедра механики композитов(Совм.)(МГУ) Ведущий научный сотрудник, Профессор
0.821	д.ф.-м.н. член-корр. Ломакин Евгений Викторович	Кафедра теории пластичности(МГУ) Заведующий кафедрой
0.721	д.ф.-м.н. доц. Никабадзе Михаил Ушангиевич	Кафедра механики композитов(МГУ) Профессор
0.707	д.ф.-м.н. проф. Тарлаковский Дмитрий Валентинович	202 Лаборатория динамических испытаний(МГУ) Заведующий лабораторией
0.547	д.ф.-м.н. Горбачев Владимир Иванович	205 Лаборатория прочности и ползучести при высоких температурах(Совм.)(МГУ); Кафедра механики композитов(МГУ) Ведущий научный сотрудник, Заведующий кафедрой
0.462	д.ф.-м.н. проф. Васин Рудольф Алексеевич	None
0.458	д.ф.-м.н. доц. Вершинин Анатолий Викторович	Кафедра вычислительной механики(МГУ) Доцент

Рис. 2. Результат поиска по запросу «Теория упругости»

Show by 10 items Search:

Тип работы	Название работы	Релевантность
Публикация	Автоматизированная система тематического анализа информации	13
Публикация	Методы и средства тематического анализа данных в больших системах на основе ключевых слов и косвенных связей между ними	13
Доклад	Методы и средства тематического анализа данных в больших системах на основе ключевых слов и косвенных связей между ними	13
Публикация	Интеллектуальная Система Тематического Исследования НАучно-технической информации (ИСТИНА)	10
Достежение	Развитие и поддержка ИАС «Наука МГУ» («ИСТИНА»)	10
Проект	Развитие и сопровождение системы подготовки принятия решений на основе анализа информации о результатах научно-исследовательской, педагогической и инновационной деятельности ИАС «ИСТИНА»	10
Проект	Математическое и программное обеспечение сложных информационно-вычислительных систем	10

Рис. 3. Список работ сотрудника по заданной запросом тематике

Автоматический поиск экспертов предполагается использовать для подбора рецензентов в журналы. Кроме того, планируется предоставить пользователям возможность поиска экспертов по ключевым словам на основной странице поиска системы.

ЗАКЛЮЧЕНИЕ

В настоящий момент сервисы тематического поиска экспертов в русскоязычном сегменте интернета представлены недостаточно широко. Использование алгоритмов тематического анализа данных наукометрических систем позволяет создавать качественные сервисы для решения задачи поиска экспертов. В алгоритме для тематического поиска и ранжирования, представленном в настоящей работе, используется комбинирование полнотекстовых и графовых методов анализа. Программная реализация алгоритма апробирована на данных наукометрической системы ИАС ИСТИНА.

СПИСОК ЛИТЕРАТУРЫ

1. Группа «Помогите журналисту».
URL: <http://www.facebook.com/groups/pomogitej>
2. Группа «Герои и эксперты для СМИ».
URL: <http://www.facebook.com/groups/1792146334361293>
3. Биржа копирайтинга. URL: <http://www.textsale.ru>
4. Биржа копирайтинга Адвего. URL: <http://Advego.ru>
5. Биржа контента eTXT. URL: <http://eTXT.ru>
6. Поисковая система Search Engine to Find Experts from Universities.
URL: <http://network.expertisefinder.com>
7. Поисковая система CiteSerX. URL: <http://citeseerx.ist.psu.edu>
8. Поисковая система Semantic Scholar. URL: <http://www.semanticscholar.org>
9. Поисковая система ScienceDirect. URL: <http://www.sciencedirect.com>
10. Издательство Шпрингер. URL: <http://link.springer.com>
11. Поисковая система издательства IEEE. URL: <http://ieeexplore.ieee.org>
12. Поисковая система Digital Library. URL: <http://dl.acm.org>
13. Система цитирования Google Scholar. URL: <http://scholar.google.com>
14. Система поиска авторов Microsoft Academic.
URL: <http://academic.microsoft.com>
15. Система поиска авторов Aminer. URL: <http://www.aminer.cn>
16. *Vasenin V., Zanchurin M., Kozitsyn A. et al.* Architectural and technological aspects of the cloud data analysis system development, case of istina system // 5th International Conference on Actual Problems of System and Software Engineering,

APSSE 2017. Vol. 1989 of CEUR Workshop Proceedings (CEUR-WS.org). CEUR Workshop Proceedings. 2017. Vol. 1989. P. 90–96.

17. Описание статьи «Architectural and technological aspects of the cloud data analysis system development, case of Istina system» в системе Microsoft Academic. URL: <http://academic.microsoft.com/paper/2771441762>

18. *Васенин В.А., Афонин С.А., Зензинов А.А. и др.* Механизмы системы ИСТИНА для интеллектуального анализа состояния и стимулирования хода выполнения проектов в сфере науки и высшего образования // Научный сервис в сети Интернет: труды XXI Всероссийской научной конференции (23–28 сентября 2019 г., г. Новороссийск) / Под ред. В.В. Воеводина. ИПМ им. М.В. Келдыша. Москва, 2019. С. 210–221.

19. Описание статьи «Mechanisms of <ISTINA> system for intelligent state analysis and progress stimulation for projects in the sphere of science and higher education» в системе Microsoft Academic.

URL: <http://academic.microsoft.com/paper/2986189703>

20. Система цитирования РИНЦ. URL: <http://elibrary.ru>

21. *Vasenin Valery, Lunev Kirill, Afonin Sergey, and Shachnev Dmitry.* Methods for intelligent data analysis based on keywords and implicit relations: The case of “Istina” data analysis system. In Proc. of the International Conference Actual Problems of Systems and Software Engineering (APSSE 2019), IEEE Conference Proceedings United States. 2019. P. 151–155.

22. *Васенин В.А.* Архитектурно-технологические аспекты разработки и сопровождения больших информационно-аналитических систем в сфере науки и образования // Программная инженерия. 2017. Т. 8, № 10. С. 448–455.

23. *Афонин С.А., Голомазов Д.Д., Козицын А.С.* Использование систем семантического анализа для организации поиска научно-технической информации // Программная инженерия. 2012. № 2. С. 29–34.

24. *Козицын А.С., Афонин С.А.* Разрешение неоднозначностей при определении авторов публикации с использованием графов соавторства в больших коллекциях библиографических данных // Программная инженерия. 2017. №8(12). С. 556–562.

25. *Козицын А.С., Афонин С.А.* Алгоритм разрешения неоднозначности имен

авторов в ИАС ИСТИНА // Современные информационные технологии и ИТ-образование. 2020. Т. 16, №1. С. 108–117.

26. *Shachnev D.A.* Searching for activity results and experts in a given subject area, taking results significance into account // Программная инженерия. 2021. Vol. 12, no. 5. P. 260–266.

27. *Козицын А.С., Афонин С.А., Шачнев Д.А.* Метод оценки тематической близости научных журналов // Программная инженерия. 2020. №6. С. 335–341.

28. *David F. Gleich.* PageRank Beyond the Web // Proc. Society for Industrial and Applied Mathematics. 2015. P. 321–363.

METHOD FOR EXPERT SEARCH USING SCIENTOMETRIC SYSTEM DATA

A. S. Kozitsin¹ [0000-0002-8065-9061], **S. A. Afonin**² [0000-0003-3058-9269]

^{1,2} Institute of Mechanics Lomonosov Moscow State University

¹alexanderkz@mail.ru, ²serg@msu.ru

Abstract

The use of modern methods of thematic analysis for the analytical processing of information is currently used in almost all areas of human activity, including scientometrics. Many scientometric and citation systems, including the world famous WoS, Scopus, Google Scholar, develop thematic categories for searching and processing information. Most important tasks that can be solved using thematic classification methods are: assessment of the dynamics of the development of thematic areas in the organization, country and in world science; search for articles on a given topic; search and assessment of the authority of experts; search for journal for publication and other relevant tasks. The Lomonosov Moscow State University is currently developing and using the system ISTINA. In this project, algorithms have been created that solve some of the problems listed. Scientific research is underway to create new effective mathematical models and algorithms in this area.

Keywords: *thematic search, bibliographic data, expert search, information systems, scientometrics*

REFERENCES

1. Gruppa «Pomogite zhurnalistu».
URL: <http://www.facebook.com/groups/pomogitej>
2. Gruppa «Geroi i eksperty dlia SMI».
URL: <http://www.facebook.com/groups/1792146334361293>
3. Birzha kopiraitinga. URL: <http://www.textsale.ru>
4. Birzha kopiraitinga Advego. URL: <http://Advego.ru>
5. Birzha kontenta eTXT. URL: <http://eTXT.ru>
6. Poiskovaia sistema Search Engine to Find Experts from Universities.
URL: <http://network.expertisefinder.com>
7. Poiskovaia sistema CiteSerX. URL: <http://citeseerx.ist.psu.edu>
8. Poiskovaia sistema Semantic Scholar.
URL: <http://www.semanticscholar.org>
9. Poiskovaia sistema ScienceDirect. URL: <http://www.sciencedirect.com>
10. Izdatelstvo Springer. URL: <http://link.springer.com>
11. Poiskovaia sistema izdatelstva IEEE. URL: <http://ieeexplore.ieee.org>
12. Poiskovaia sistema Digital Library. URL: <http://dl.acm.org>
13. Sistema tsitirovaniia Google Scholar. URL: <http://scholar.google.com>
14. Sistema poiska avtorov Microsoft Academic.
URL: <http://academic.microsoft.com>
15. Sistema poiska avtorov Aminer. URL: <http://www.aminer.cn>
16. *Vasenin V., Zanchurin M., Kozitsyn A. et al.* Architectural and technological aspects of the cloud data analysis system development, case of istina system // 5th International Conference on Actual Problems of System and Software Engineering, APSSE 2017. Vol. 1989 of CEUR Workshop Proceedings (CEUR-WS.org). CEUR Workshop Proceedings. 2017. Vol. 1989. P. 90–96.
17. Opisanie stati “Architectural and technological aspects of the cloud data analysis system development, case of Istina system” v sisteme Microsoft Academic.
URL: <http://academic.microsoft.com/paper/2771441762>
18. *Vasenin V.A., Afonin S.A., Zenzinov A.A. i dr.* Mekhanizmy sistemy ISTINA dlia intellektualnogo analiza sostoianiia i stimulirovaniia khoda vypolneniia proektov v

sfere nauki i vysshego obrazovaniia // Nauchnyi servis v seti Internet: trudy XXI Vse-rossiiskoi nauchnoi konferentsii (23–28 sentiabria 2019 g., g. Novorossiisk) / Pod red. V.V. Voevodin. IPM im. M.V. Keldysha. Moskva, 2019. S. 210–221.

19. Opisanie stati “Mechanisms of <ISTINA> system for intelligent state analysis and progress stimulation for projects in the sphere of science and higher education” v sisteme Microsoft Academic.

URL: <http://academic.microsoft.com/paper/2986189703>

20. Sistema tsitirovaniia RINTs. URL: <http://elibrary.ru>

21. *Vasenin Valery, Lunev Kirill, Afonin Sergey, and Shachnev Dmitry.* Methods for intelligent data analysis based on keywords and implicit relations: The case of “istina” data analysis system. In Proc. of the International Conference Actual Problems of Systems and Software Engineering (APSSE 2019), IEEE Conference Proceedings United States, 2019. P. 151–155.

22. *Vasenin V.A.* Arkhitekturno-tekhnologicheskie aspekty razrabotki i so-provozhdeniia bolshikh informatsionno-analiticheskikh sistem v sfere nauki i obra-zovaniia // Programmnaia inzheneriia. 2017. T. 8, № 10. S. 448–455.

23. *Afonin S.A., Golomazov D.D., Kozitsyn A.S.* Ispolzovanie sistem seman-ticheskogo analiza dlia organizatsii poiska nauchno-tekhnicheskoi informatsii // Pro-grammnaia inzheneriia. 2012. № 2. S. 29–34.

24. *Kozitsyn A.S., Afonin S.A.* Razreshenie neodnoznachnostei pri opredelenii avtorov publikatsii s ispolzovaniem grafov soavtorstva v bolshikh kolleksiakh biblio-graficheskikh dannykh // Programmnaia inzheneriia. 2017. №8(12). S. 556–562.

25. *Kozitsyn A.S., Afonin S.A.* Algoritm razresheniia neodnoznachnosti imen avtorov v IAS ISTINA // Sovremennye informatsionnye tekhnologii i IT-obrazovanie. 2020. T. 16, №1. S. 108–117.

26. *Shachnev D.A.* Searching for activity results and experts in a given subject area, taking results significance into account // Programmnaia Inzheneriia. 2021. Vol. 12, no. 5. P. 260–266.

27. *Kozitsyn A.S., Afonin S.A., Shachnev D.A.* Metod otsenki tematicheskoi blizosti nauchnykh zhurnalov // Programmnaia Inzheneriia. 2020. №6. S. 335–341.

28. *David F. Gleich.* PageRank Beyond the Web // Proc. Society for Industrial and Applied Mathematics. 2015. P. 321–363.

СВЕДЕНИЯ ОБ АВТОРАХ



КОЗИЦЫН Александр Сергеевич – ведущий научный сотрудник, к. ф.-м. н., окончил мехмат МГУ им. М.В. Ломоносова. Специалист в области информационного поиска и баз данных.

Alexander Sergeevich KOZITSYN – Leading Researcher, Ph. D., graduated from M.V. Lomonosov Moscow State University. Specialist in the field of information retrieval and database.

email: alexanderkz@mail.ru,
ORCID: 0000-0002-8065-9061



АФОНИН Сергей Александрович – ведущий научный сотрудник, к. ф.-м. н., окончил мехмат МГУ им. М.В. Ломоносова. Специалист в области регулярных языков и информационных систем.

Sergey Alexandrovich AFONIN – Leading Researcher, Ph. D., graduated from M.V. Lomonosov Moscow State University. Specialist in the field of regular languages and information systems.

email: serg@msu.ru,
ORCID:0000-0003-3058-9269

Материал поступил в редакцию 21 октября 2021 года

УДК 004.946

ТЕХНОЛОГИЯ СИТУАЦИОННОГО МОДЕЛИРОВАНИЯ В СИСТЕМАХ ВИРТУАЛЬНОГО ОКРУЖЕНИЯ

М. В. Михайлюк^[0000-0002-7793-080X], Д. А. Кононов^[0000-0002-6059-5590],

Д. М. Логинов^[0000-0002-2717-6909]

*Научно-исследовательский институт системных исследований
Российской академии наук, г. Москва*

mix@niisi.ras.ru

Аннотация

Обсуждена технология моделирования различных ситуаций в системах виртуального окружения, которые являются компьютерными трехмерными моделями реальной или искусственной среды. Пользователь может рассматривать эти сцены непосредственно на экране компьютера, настенном экране, в стерео очках, в очках виртуальной реальности и т. д. Он также может перемещаться внутри виртуальной сцены и взаимодействовать с ее объектами. В свою очередь среда также может изменяться. Это позволяет проводить в системе виртуального окружения моделирование различных ситуаций (ситуационное моделирование). При таком моделировании задается некоторая статическая или динамическая обстановка в системе виртуального окружения, в которой оператор должен выполнить поставленные перед ним задачи.

Предложен механизм задания ситуаций путем изменения виртуальной трехмерной сцены с помощью конфигурационных файлов и виртуальных пультов управления. Для записи конфигурационных файлов разработан специальный язык, а для создания виртуальных пультов управления – специальный редактор. Представлена апробация предложенных методов на примере двух виртуальных сцен: полигона для мобильных роботов и реактивного рюкзака спасения космонавта в открытом космосе.

Ключевые слова: *открытый доступ, система виртуального окружения, ситуационное моделирование, трехмерная сцена, конфигурационный файл, виртуальный пульт управления.*

ВВЕДЕНИЕ

Системы виртуального окружения (СВО) представляют собой компьютерные трехмерные модели реальной или искусственной среды, которые пользователь может наблюдать непосредственно на экране компьютера, настенном экране, в стерео очках или в очках виртуальной реальности и т. д. Кроме наблюдения, пользователь может перемещаться внутри этой среды, взаимодействовать с ее объектами и анализировать результаты этого взаимодействия. Среда может также изменяться самостоятельно или под действием каких-либо факторов. Все это приводит к возможности проведения моделирования в СВО.

Известно несколько типов моделирования, которые могут использоваться в СВО. При имитационном моделировании создается компьютерная модель реальных или разрабатываемых системы или процесса, на которых изучаются результаты моделирования (поведение и параметры системы или процесса) или предсказываются будущие результаты. Существенными здесь являются визуализация всех процессов и их визуальное наблюдение. Сценарное моделирование позволяет изучать стойкость и живучесть сложных систем. Его задачами являются недопущение нештатных ситуаций, управление рисками, восстановление живучести систем и др. Структура системы и взаимодействие элементов при ее функционировании представляют в виде ориентированного графа, дугам и вершинам которого присваивают параметры и функционалы, адекватно описывающие процессы функционирования элементов исследуемой (моделируемой) системы. Визуализация такой системы в СВО позволяет наглядно наблюдать происходящие процессы. Имитационно-тренажерные комплексы предназначены для профессиональной подготовки операторов путем многократного выполнения ими необходимых действий. В этих комплексах СВО используется для моделирования динамики и визуализации окружающей обстановки. Расширение задач таких комплексов путем создания различных ситуаций для операторов приводит к ситуационному моделированию.

Ситуационное моделирование [1] создает особую статическую или динамическую обстановку в СВО, в которой оператор должен выполнить поставленные задачи. При этом он должен следовать как объективным свойствам созданной ситуации, так и своим субъективным представлениям о том, как действовать в такой

ситуации. Здесь важны не столько обучение конкретному ролевому поведению, сколько умение справляться со сложными ситуациями. Инструктор может предлагать как отдельные ситуации, так и несколько ситуаций, следующих друг за другом. Такое моделирование позволяет проверить не только квалификацию оператора, но и его психологические качества: смелость, склонность к риску, настойчивость, обучаемость, эмоциональность, стрессоустойчивость, адекватность, целенаправленность, потенциал работоспособности и др. В работе [2] рассмотрена ситуационная мотивация в решении такой проблемы, как состояние ситуативной когнитивной и эпистемической готовности к решению проблемы для уменьшения воспринимаемого несоответствия между ожидаемым и реальным состояниями. В [3] предложена классификация ситуаций и соответствующих им алгоритмов поведения на примере управления малым турбореактивным двигателем МРМ. В статье [4] рассмотрена работа оператора с пультом управления в виртуальном интерьере космического модуля «Пирс» для осуществления выхода космонавта в открытый космос. Целый комплекс возможных задач представлен в работе [5]. В ней, на основе анализа подготовки космонавтов к внекорабельной деятельности выделены классы задач, которые физически сложно моделируются на технических средствах. К ним относятся задачи по аварийным ситуациям, управлению перспективными средствами, информационного обеспечения космонавтов и др. Некоторые другие задачи представлены в работах [6–8].

В данной работе рассмотрена задача ситуационного моделирования в системах виртуального окружения на примере двух виртуальных сцен: полигона для мобильных роботов и реактивного рюкзака спасения космонавта в открытом космосе. Предложены методы задания различных ситуаций в одной 3D виртуальной сцене с помощью конфигурационных файлов и виртуальных пультов управления, а также возможные задачи ситуационного моделирования в СВО.

1. СИСТЕМА ВИРТУАЛЬНОГО ОКРУЖЕНИЯ

Система виртуального окружения состоит из трехмерной (3D) виртуальной сцены, подсистемы управления, подсистемы динамики и подсистемы визуализации (см. рис. 1). Виртуальная 3D сцена создается в системе трехмерного моделирования (например, 3ds Max) и задает сцену, в которой будут происходить все действия. На рис. 2 слева показана сцена полигона мобильных роботов, а справа

– сцена космонавта с реактивным ранцем спасения в открытом космосе. Выбранная сцена на начальном этапе загружается в подсистемы динамики и визуализации. Подсистема управления включает пульта управления динамическими объектами (например, роботами или реактивными двигателями ранца), функциональные схемы и программные модули вычисления управляющих сигналов, являющихся результатом воздействия оператора на элементы пульта управления. При этом можно использовать как реальные, так и виртуальные пульта управления.

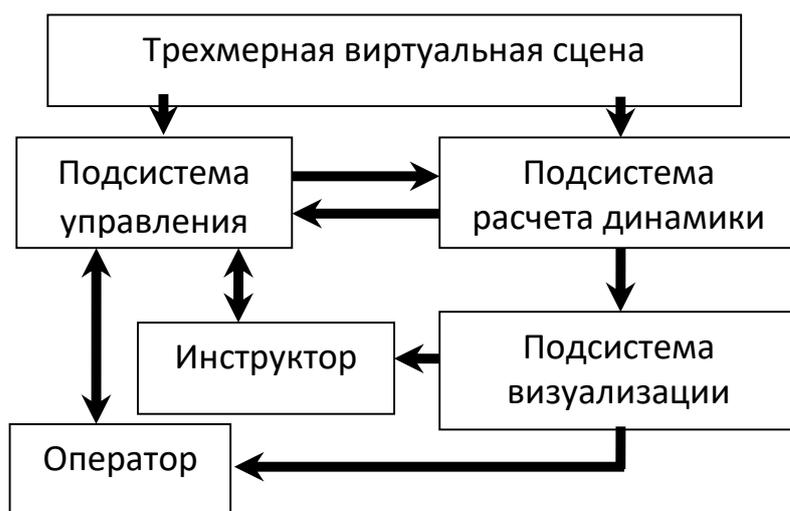


Рис. 1. Система виртуального окружения.

Для создания виртуальных пультов управления разработан специальный интерактивный редактор, содержащий большой набор типов управляющих элементов (кнопки, тумблеры, джойстики, регуляторы и т. д.). Функциональные схемы создаются в другом специальном редакторе, который включает широкий набор функциональных блоков различных функций (арифметических, логических, тригонометрических, динамических, автоматического управления, генераторов сигналов и др.). Каждый блок имеет входы и выходы, которые можно соединять в редакторе линиями, получая таким образом функциональную схему. На вход схемы подаются управляющие воздействия оператора, показания датчиков из виртуальной сцены, параметры настройки СВО и др. Вычисленные сигналы передаются в подсистему динамики, которая рассчитывает новые координаты, углы ориентации и состояния управляемых объектов через промежуток Δt времени моделирования. При этом учитываются все динамические параметры, столкнове-

ния (коллизии) объектов, силы трения, гравитация и т. д. Результаты расчетов передаются в подсистему визуализации. Эта подсистема осуществляет синтез изображения виртуальной сцены с новыми параметрами динамических объектов. Также моделируются различные типы освещения, специальные объекты (огонь, струи воды и пены, рельеф и др.), а также состояния окружающей среды (время суток, дождь, снег, туман и т. д.). Так как один цикл работы всех подсистем занимает не более 40 миллисекунд, то у оператора создается впечатление непрерывного управляемого процесса в системе виртуального окружения.



Рис. 2. Трехмерная сцена в системе виртуального окружения.

В ФГУ ФНЦ НИИСИ РАН разработана оригинальная СВО VirSim, включающая все описанные подсистемы. Ее основное назначение заключается в использовании в качестве тренажера операторов управления сложными динамическими системами. Поэтому работа всех подсистем должна быть максимально реалистичной, чтобы не формировать у операторов так называемые ложные навыки. Эта система использовалась для подготовки операторов мобильных и антропоморфных роботов, а также для тренировки космонавтов на МКС. На рис. 2 показаны примеры виртуальных сцен в этой СВО.

Сцену СВО можно рассматривать как некоторую ситуацию (ситуационную модель), определяемую взаимным расположением объектов и их свойствами.

2. СИТУАЦИОННОЕ МОДЕЛИРОВАНИЕ

Под ситуационным моделированием здесь понимаются создание некоторой ситуации (статической или динамической) в СВО и решение в ней оператором (или группой операторов) некоторой задачи. Задание ситуации, постановка за-

дачи, контроль и оценка действий оператора осуществляются инструктором. Инструктор может предлагать как отдельные ситуации, так и несколько ситуаций, идущих подряд. Целью моделирования является не только обучение оператора выполнению известных действий, но и умению найти правильное решение в сложных ситуациях. Поэтому оператор не только должен учитывать объективные свойства заданной ситуации, но и использовать свои способности действовать в сложной и неожиданной обстановке. Примером являются нештатные ситуации в моделируемой системе. По окончании тренировки обычно проводятся анализ и самоанализ участниками процесса своих действий и их результатов. Это позволяет приобрести опыт действия в таких ситуациях, не подвергая опасности свою жизнь и здоровье. В работе [5] выделены классы задач в процессе подготовки космонавтов к внекорабельной деятельности, которые физически сложно смоделировать на обычных технических средствах, но возможно с использованием технологий виртуальной реальности. К первому классу относятся задачи по информационному обеспечению космонавта, а именно, демонстрация сложившейся ситуации и необходимых действий космонавта в этой ситуации. Ко второму классу относится моделирование действий космонавта по выполнению ряда задач и демонстрации результатов допущенных ошибок (незакрепления страховочного фала, неправильной траектории отброса запускаемых малых спутников, пропуск аварийного сигнала и др.). Третий класс включает задачи по управлению перспективными средствами перемещения космонавта (спасательный ранец, робототехнические системы, мобильные роботы и т. д.). Аналогичную классификацию можно провести для любой области деятельности.

3. МЕТОДЫ ЗАДАНИЯ СИТУАЦИИ В СВО

Статическая ситуация — это ситуация, в которой все объекты и окружающая обстановка не изменяют своих параметров в процессе моделирования. Самый простой способ моделирования статической ситуации — это создание для СВО новой трехмерной сцены в системе 3D моделирования. Например, для этой цели можно использовать систему 3ds Max. Однако во многих задачах для различных ситуаций в качестве базовой может использоваться одна и та же сцена. В этом случае возникнут многочисленные ее конкретные модификации, которые будут

занимать большой объем памяти и мало отличаться друг от друга. В случае изменения базовой сцены придется отслеживать изменения всех ее модификаций.



Рис. 3. Изменение ситуации с помощью конфигурационного файла.

Другой способ задания ситуации возможен с помощью конфигурационного файла. Конфигурационный файл представляет собой xml файл, в котором для любого объекта можно задать значение любого его параметра. Для объекта можно задать положение и ориентацию, а также сделать его невидимым или заменить на поврежденный; источник освещения можно включить или выключить, задать время суток, включить дождь или снег и задать их интенсивности и т. д. На рис. 3 слева показан пример сцены, а справа – той же сцены, в которой некоторые объекты заданы невидимыми с помощью конфигурационного файла.

Для управления конфигурационными файлами можно использовать виртуальные пульты управления. Такие пульта создаются с помощью специально разработанного редактора, который позволяет разместить на подложке пульта элементы управления (кнопки, тумблеры, переключатели и т. д.). Положения элементов управления при моделировании передаются в функциональную схему, которая строится из функциональных блоков различного типа (арифметических, тригонометрических, управляющих и др.). В частности, в ней имеются блоки отслеживания времени и блоки запуска на исполнение конфигурационных файлов.

С помощью этого механизма можно изменить ситуацию в виртуальной сцене в определенный момент времени. На рис. 4 показан пример пульта инструктора, с помощью которого он может задать одну из 12 ситуаций и выбрать погодные условия (дождь или снег).

4. ВОЗМОЖНЫЕ ЗАДАЧИ СИТУАЦИОННОГО МОДЕЛИРОВАНИЯ

Рассмотрим возможные задачи ситуационного моделирования на примере двух сцен СВО: полигона мобильных роботов и спасательного реактивного ранца (сейфера) космонавта. В первой сцене (см. рис. 2 слева) оператор с помощью виртуального или реального пульта управляет движением робота и может осуществлять захват объектов манипулятором. Во второй сцене (см. рис. 2 справа) оператор с помощью специального реального или виртуального джойстика управляет реактивным ранцем, перемещая космонавта в пространстве.



Рис. 4. Ситуационный пульт управления.

На начальном этапе оператор должен освоить управление и выполнение простых базовых операций (изучение пульта управления, перемещение управляемого средства вперед и назад, поворот влево и вправо, захваты объектов и т. д.).

Более сложные задачи состоят в перемещении робота (или сейфера) в заданную точку. Точка может задаваться координатами (при этом необходимо моделировать систему локального позиционирования и выводить текущие координаты на пульт управления) или описанием (например, вход в ангар, выходной люк Международной космической станции и др.). К такому типу задач относятся также поиски заданных объектов или инспекция окружающей обстановки. При поиске объектов могут использоваться специальные виртуальные датчики. Например, виртуальный γ -детектор показывает уровень γ -излучения радиоактивного объекта, попадающего в его конус действия. Его можно использовать для поиска зараженных радиоактивных объектов с целью их дальнейшей транспортировки в

специализированные контейнеры. Объект может быть также задан описанием (например, источник огня, бочка с топливом, солнечная батарея, определенный космический модуль и т. д.). Задачи инспекции окружающей обстановки могут включать осуществление съемки виртуальной камерой (включая панорамную съемку 360 градусов), поиск повреждений, проверку правильности креплений конструкций и т. д.). Выполнение всех этих операций можно ограничивать временем и наличием препятствий.



Рис. 5. Задачи ситуационного моделирования.

Более сложными операциями являются тушение горящих объектов с помощью струи воды или пены из водомета мобильного робота, ориентация при выходе из строя какого-либо датчика, «застревание» робота в яме, восстановление его работоспособности и др. В сцене с сейфером можно отрабатывать автоматический возврат (в случае запотевания скафандра или потери видимости МКС), возврат к ближайшему поручню (в случае недостатка топлива для перелета к шлюзу), выход из строя одного двигателя, стабилизацию космонавта (в случае его закручивания) и т. д. Имея в сцене подстилающую поверхность Земли, можно изменять время суток, ставить задачи распознавания места пролета, выявление пожаров, наводнений и других катаклизмов. На рис. 5 слева показан пример тушения роботом очага возгорания, при этом оператор управляет процессом с помощью виртуального пульта управления в левом нижнем углу. Справа на этом же рисунке показана модель подстилающей земной поверхности с высоты 300 км, на которой необходимо распознать видимый участок местности.

ЗАКЛЮЧЕНИЕ

В работе предложены методы задания различных ситуаций, а также возможные задачи ситуационного моделирования в СВО на примере двух виртуальных сцен: полигона для мобильных роботов и реактивного рюкзака спасения космонавта в открытом космосе. При этом ситуационное моделирование осуществлялось в СВО VirSim, разработанной в ФГУ ФНЦ НИИСИ РАН. Апробация создания различных ситуаций с помощью 3D виртуальных сцен, конфигурационных файлов и виртуальных пультов управления показала адекватность предложенных методов поставленным задачам.

Благодарности. Публикация выполнена в рамках государственного задания ФГУ ФНЦ НИИСИ РАН «Проведение фундаментальных научных исследований (47 ГП)» по теме № FNEF-2021-0012 «Системы виртуального окружения: технологии, методы и алгоритмы математического моделирования и визуализации. 0580-2021-0012», Рег. № 121031300061-2.

СПИСОК ЛИТЕРАТУРЫ

1. *Поспелов Д.А.* Ситуационное управление: теория и практика. М.: Наука, Гл. ред. физ.-мат. лит., 1986, 288 с.
2. *Jeong-Nam Kim, James E. Grunig.* Problem Solving and Communicative Action: A Situational Theory of Problem Solving // Journal of Communication. 2011. V. 61. P. 120–149.
3. *Andoga R., Főző L., Madarász L.* Digital Electronic Control of a Small Turbojet Engine MPM 20 // Acta Polytechnica Hungarica. 2007. V. 4. No. 4. P. 83–95.
4. *Maltsev A.V., Mikhaylyuk M.V.* Virtual Environment System for Pirs Space Module Interior // CEUR Workshop Proceedings: Proc. of the 29th International Conference on Computer Graphics and Vision. 2019. V. 2485.
URL: <http://ceur-ws.org/Vol-2485/paper1.pdf>
5. *Алтунин А.А., Долгов П.П., Жамалетдинов Н.Р., Иродов Е.Ю., Коренной В.С.* Направления применения технологий виртуальной реальности при подготовке космонавтов к внекорабельной деятельности // Пилотируемые полеты в космос. 2021. № 1 (38). С. 72–88.

6. Maltsev A.V., Mikhaylyuk M.V. Visualization and virtual environment technologies in the tasks of cosmonaut training // *Scientific Visualization*. 2020. V. 12, No. 3. P. 16–25.

7. Михайлюк М.В., Мальцев А.В., Тимохин П.Ю., Страшнов Е.В., Крючков Б.И., Усов В.М. Системы виртуального окружения для прототипирования на моделирующих стендах использования космических роботов в пилотируемых полетах // *Пилотируемые полеты в космос*. 2020. № 2 (35). С. 61–75.

8. Tomchinskaya T., Shaposhnikova M., Dudakov N. Training Beginners and Experienced Drivers using mobile-based Virtual and Augmented Reality // *CEUR Workshop Proceedings: Proc. of the 30th International Conference on Computer Graphics and Vision*. 2020. V. 2744. URL: <http://ceur-ws.org/Vol-2744/paper69.pdf>

SITUATIONAL MODELING TECHNOLOGY IN VIRTUAL ENVIRONMENT SYSTEMS

M. V. Mikhaylyuk^[0000-0002-7793-080X], D. A. Kononov^[0000-0002-6059-5590],

D. M. Loginov^[0000-0002-2717-6909]

Scientific Research Institute for System Analysis RAS, Moscow

mix@niisi.ras.ru

Abstract

The technology of modelling various situations in virtual environment systems, which are computer three-dimensional models of a real or artificial environment, is discussed. The user can view these scenes directly on the computer screen, wall screen, in a stereo glasses, virtual reality glasses, etc. He can also move inside a virtual scene and interact with its objects. In turn, the environment can also change. This allows modelling of various situations (situation modelling) in the virtual environment system. With such modelling, some static or dynamic situation is set in the virtual environment system in which the operator must perform the tasks assigned to him.

A mechanism for setting situations by changing a virtual three-dimensional scene using configuration files and virtual control panels is proposed. A special language has been developed for writing configuration files, and a special editor has been developed

for creating virtual control panels. The approbation of the proposed methods is presented on the examples of two virtual scenes: a training ground for mobile robots and a jet backpack for the rescue of an astronaut in outer space.

Keywords: *virtual environment system, situational modeling, three-dimensional scene, configuration file, virtual control panels*

REFERENCES

1. *Pospelov D.A.* Situatsionnoe upravlenie: teoriia i praktika. M.: Nauka, Gl. red. fiz.-mat. lit., 1986. 288 s.
2. *Jeong-Nam Kim, James E. Grunig.* Problem Solving and Communicative Action: A Situational Theory of Problem Solving // *Journal of Communication.* 2011. V. 61. P. 120–149.
3. *Andoga R., Főző L., Madarász L.* Digital Electronic Control of a Small Turbojet Engine MPM 20. *Acta Polytechnica Hungarica.* 2007. V. 4, No. 4. P. 83–95.
4. *Maltsev A.V., Mikhaylyuk M.V.* Virtual Environment System for Pirs Space Module Interior // *CEUR Workshop Proceedings: Proc. of the 29th International Conference on Computer Graphics and Vision.* 2019. V. 2485. URL: <http://ceur-ws.org/Vol-2485/paper1.pdf>
5. *Altunin A.A., Dolgov P.P., ZHamaletdinov N.R., Irodov E.YU., Korennoj V.S.* Napravleniya primeneniya tekhnologij virtual'noj real'nosti pri podgotovke kosmonavtov k vnekorabel'noj deyatel'nosti // *Pilotiruemye polety v kosmos.* 2021. No. 1 (38). S. 72–88.
6. *Maltsev A.V., Mikhaylyuk M.V.* Visualization and virtual environment technologies in the tasks of cosmonaut training // *Scientific Visualization.* 2020. V. 12, No. 3. P. 16–25.
7. *Mikhailiuk M.V., Maltsev A.V., Timokhin P.Iu., Strashnov E.V., Kriuchkov B.I., Usov V.M.* Sistemy virtualnogo okruzheniia dlia prototipirovaniia na modeliruiushchikh stendakh ispolzovaniia kosmicheskikh robotov v pilotiruemykh poletakh // *Pilotiruemye polety v kosmos.* 2020. № 2 (35). S. 61–75.
8. *Tomchinskaya T., Shaposhnikova M., Dudakov N.* Training Beginners and Experienced Drivers using mobile-based Virtual and Augmented Reality // *CEUR Workshop Proceedings: Proc. of the 30th International Conference on Computer Graphics and Vision.* 2020. V. 2744. URL: <http://ceur-ws.org/Vol-2744/paper69.pdf>

СВЕДЕНИЯ ОБ АВТОРАХ



МИХАЙЛЮК Михаил Васильевич – доктор физико-математических наук, профессор, главный научный сотрудник НИИ системных исследований РАН.

Mikhail Vasilievich MIKHAYLYUK – Doctor of Physics and Mathematics, Professor, Scientific Research Institute for System Analysis RAS, chief researcher

email: mix@niisi.ras.ru

ORCID: 0000-0002-7793-080X



КОНОНОВ Дмитрий Алексеевич – доктор технических наук, доцент, ведущий научный сотрудник НИИ системных исследований РАН.

Dmitry Alekseevich KONONOV – Doctor of Technical Sciences, Associate Professor, Scientific Research Institute for System Analysis RAS, chief researcher

email: mix@niisi.ras.ru

ORCID: 0000-0002-6059-5590



ЛОГИНОВ Дмитрий Михайлович – кандидат технических наук, старший научный сотрудник НИИ системных исследований РАН.

Dmitry Mikhailovich LOGINOV – Candidate of Technical Sciences, Scientific Research Institute for System Analysis RAS, senior researcher

email: mix@niisi.ras.ru

ORCID: 0000-0002-2717-6909

Материал поступил в редакцию 19 октября 2021 года

ОПЫТ ВЕРИФИКАЦИИ РЕАЛИЗАЦИЙ ПРОТОКОЛА TLS 1.3

А. В. Никешин¹ [0000-0001-5781-9736], **В. З. Шнитман**² [0000-0002-1509-0972]

^{1, 2}Институт системного программирования им. В.П. Иванникова РАН,
ул. А. Солженицына, 25, г. Москва, 109004

¹alexn@ispras.ru, ²vzs@ispras.ru

Аннотация

Представлен опыт верификации реализаций сервера криптографического протокола TLS версии 1.3. TLS – широко распространенный криптографический протокол, предназначенный для создания защищенных каналов передачи данных и обеспечивающий необходимую для этого функциональность: конфиденциальность передаваемых данных, целостность данных, аутентификацию сторон. Новая версия протокола TLS 1.3 была представлена в августе 2018 года и имеет ряд существенных отличий по сравнению с предыдущей версией 1.2. Ряд разработчиков протокола TLS уже включил поддержку последней версии в свои реализации. Данные обстоятельства делают актуальным проведение исследований в области верификации и безопасности реализаций новой версии протокола TLS. В работе использован новый тестовый набор для верификации реализаций протокола TLS 1.3 на соответствие спецификациям интернета, разработанный на основе спецификации RFC 8446 с использованием технологии UniTESK и методов мутационного тестирования. Текущая работа является частью проекта верификации протокола TLS 1.3 и охватывает часть дополнительной функциональности и необязательных расширений протокола.

Для тестирования реализаций на соответствие формальным спецификациям применена технология UniTESK, предоставляющая средства автоматизации тестирования на основе использования конечных автоматов. Состояния тестируемой системы задают состояния автомата, а тестовые воздействия – переходы этого автомата. При выполнении перехода заданное воздействие передается на тестируемую реализацию, после чего регистрируются реакции реализации и автоматически выносятся вердикт о соответствии наблюдаемого поведения специ-

фикации. Мутационные методы тестирования используются для обнаружения нестандартного поведения тестируемой системы с помощью передачи некорректных данных. В поток обмена протокола, создаваемый в соответствии со спецификацией, вносятся некоторые изменения: либо изменяются значения полей сообщений, сформированных на основе разработанной модели протокола, либо изменяется порядок сообщений в потоке обмена. Модель протокола позволяет вносить изменения в поток данных на любом этапе сетевого обмена, что позволяет тестовому сценарию проходить через все значимые состояния протокола и в каждом таком состоянии проводить тестирование реализации в соответствии с заданной программой. На данный момент было обнаружено несколько отклонений реализаций от спецификации.

Представленный подход доказал свою эффективность в нескольких наших проектах при тестировании сетевых протоколов, обеспечив обнаружение различных отклонений от спецификации и других ошибок.

Ключевые слова: *безопасность, TLS, TLSv1.3, протоколы, тестирование, оценка устойчивости, Интернет, стандарты, формальные методы спецификации.*

ВВЕДЕНИЕ

TLS является сегодня одним из наиболее востребованных криптографических протоколов, предназначенных для создания защищенных каналов передачи данных и обеспечивающий необходимую для этого функциональность: конфиденциальность передаваемых данных, целостность данных, аутентификацию сторон. В процессе использования протокола неизбежно выявляются его недостатки, появляются новые угрозы безопасности, что приводит к дальнейшему развитию протокола. В августе 2018 года была представлена спецификация новой версии протокола TLS 1.3, в которой, по сравнению с предыдущей версией 1.2, была существенно переработана архитектура [1, 2].

Все больше разработчиков протокола TLS включает поддержку последней версии в свои реализации. Данные обстоятельства делают актуальным проведение исследований в области верификации и безопасности реализаций новой вер-

сии протокола TLS. В процессе тестирования сетевых протоколов решается несколько важных задач: проверяется функциональная совместимость различных реализаций, проверяется соответствие реализации требованиям спецификации и устойчивость реализации к нестандартным воздействиям.

В данной работе использованы наработанные нами методики по тестированию сетевых протоколов: автоматизированное тестирование на соответствие формальным спецификациям и методы мутации данных, доказавшие свою эффективность в наших предыдущих проектах при тестировании сетевых протоколов, обеспечив обнаружение различных отклонений от спецификации и других ошибок.

1. ОСНОВНАЯ ФУНКЦИОНАЛЬНОСТЬ ПРОТОКОЛА TLS ВЕРСИИ 1.3

Архитектура TLS делится на два уровня обработки сообщений:

– Протокол рукопожатия (Handshakeprotocol) является основной частью архитектуры TLS. Он выполняет аутентификацию взаимодействующих сторон, согласовывает версию протокола, криптографические режимы и параметры безопасности, устанавливает общий ключевой материал. Встроенные механизмы защиты противодействуют вмешательству третьей стороны и понижению уровня безопасности согласованного криптографического набора.

– Протокол записей (Recordprotocol) используется для инкапсуляции протоколов более высокого уровня, в том числе протокола рукопожатия. Он использует параметры, установленные протоколом рукопожатия для защиты данных, передаваемых между партнерами. Весь трафик делится на последовательность записей, каждая из которых независимо защищается с помощью согласованных криптографических ключей. Протокол записей, работает поверх надежного транспортного протокола (как правило, TCP).

Неудачное рукопожатие или другая ошибка протокола вызывают завершение соединения, которому может предшествовать сообщение уведомления.

TLS не зависит от протоколов прикладного уровня, для которых он является прозрачным, спецификация TLS не определяет схему их взаимодействия. Решение о том, как инициировать TLS-диалог и интерпретировать результаты его работы, оставляется на усмотрение разработчиков протоколов верхнего уровня.

Ниже представлено основное полное рукопожатие TLSv1.3:

C→S: ClientHello (Key Exchange phase)
+ key_share*
+ signature_algorithms*
+ psk_key_exchange_modes*
+ pre_shared_key*

C←S: ServerHello
+ key_share*
+ pre_shared_key*
{EncryptedExtensions} (Server Parameters phase)
{CertificateRequest*}
{Certificate*} (Authentication phase)
{CertificateVerify*}
{Finished}
[Application Data*]

C→S:
{Certificate*}
{CertificateVerify*}
{Finished}

[ApplicationData] <-----> [ApplicationData]

* Обозначает необязательные или зависящие от ситуации сообщения/расширения, которые посылаются не всегда.

{ } Обозначает сообщения, защищенные с помощью сеансовых ключей для обмена рукопожатия.

[] Обозначает сообщения, защищенные с помощью сеансовых ключей для прикладных данных.

Назначение сообщений и расширений подробно описано в спецификации.

Сообщения протокола рукопожатия сгруппированы в три блока:

– Обмен ключами (Key Exchange): Сообщения Client Hello, Server Hello. Устанавливает ключевой материал и выбирает криптографические параметры.

– Параметры сервера (Server Parameters): Сообщения EncryptedExtensions, CertificateRequest. Устанавливает дополнительные параметры рукопожатия (например, запрос аутентификации клиента, поддержка протокола прикладного уровня и др.).

– Аутентификация (Authentication): Сообщения Certificate, Certificate Verify, Finished. Аутентифицирует сервер (и, при необходимости, клиента), подтверждает созданные сеансовые ключи и обеспечивает целостность всего обмена рукопожатия.

2. ОСНОВНЫЕ ОСОБЕННОСТИ НОВОЙ ВЕРСИИ ПРОТОКОЛА TLS

Все сообщения, отправляемые после фазы обмена ключами (т. е. после сообщения Server Hello), зашифрованы соответствующими ключами. Это одна из особенностей данной версии протокола TLS – шифровать как можно больше передаваемых данных.

Сообщение Encrypted Extensions содержит расширения, ответные для соответствующих расширений Client Hello, которые не требуются для создания криптографических параметров. Такие расширения ранее отправлялись в открытом виде.

После формирования сообщения Server Hello сервер получает возможность вычислить все необходимые сеансовые ключи. Это позволяет начать отправку прикладных данных до приема сообщений аутентификации клиента. Однако необходимо понимать, что в этой точке любые данные посылаются не аутентифицированному клиенту.

Набор сообщений протокола рукопожатия был переработан, чтобы стать более согласованным и ликвидировать излишние сообщения, такие как Change Cipher Spec (за исключением случаев обеспечения обратной совместимости с промежуточными сетевыми устройствами (middlebox)).

Механизм согласования версий перемещен в расширения, что должно улучшить совместимость с существующими серверами, которые часто неправильно реализуют согласование версий. Также изменен механизм защиты от понижения номера версии.

Изменена концепция набора шифров, чтобы разделить механизмы аутентификации и обмена ключами от алгоритмов защиты данных. Список поддерживаемых алгоритмов переработан и избавлен от всех устаревших и ненадежных алгоритмов. Оставшиеся алгоритмы шифрования являются алгоритмами аутентифицированного шифрования со связанными данными (Authenticated Encryption with Associated Data, AEAD).

Функции вычисления ключей были переделаны. Новая конструкция упрощает специалистам анализ криптографических механизмов благодаря улучшенным свойствам разделения ключей.

Переработан механизм возобновления сеанса. Теперь для этого используется общий ключ, согласованный в предыдущем обмене рукопожатия.

Добавлен новый механизм передачи данных 0-RTT (zeroround-triptime), позволяющий несколько сократить количество обменов данными между партнерами.

Также стоит отметить, что в более ранних спецификациях TLS определенные части было трудно понять, что приводило к возникновению проблем с функциональной совместимостью реализаций и безопасностью. В данной спецификации более четко проработаны вопросы поведения реализаций в различных ситуациях.

3. РАСШИРЕННАЯ ФУНКЦИОНАЛЬНОСТЬ ПРОТОКОЛА TLS1.3

Кроме базового полного рукопожатия, TLS как расширяемый протокол предоставляет дополнительные, необязательные сервисы безопасности, согласуемые, как правило, через механизм расширений. Ниже рассмотрена часть таких сервисов, указанных в спецификации TLS.

Режим PSK

TLS1.3 позволяет использовать сокращенный режим рукопожатия с использованием заранее распределенных ключей (pre-sharedkey, PSK) [2]. Данный режим использует меньше сообщений для обмена и меньше трудоемких вычислений с ключами и может быть полезен в небольших системах, в которых проще сконфигурировать PSK на каждом узле, чем разворачивать инфраструктуру для

работы с сертификатами. Также стороны могут иметь свой механизм установления общего секретного ключа.

C→S: ClientHello
+ key_share*
+ psk_key_exchange_modes

+ pre_shared_key

C←S: ServerHello
+ pre_shared_key
+ key_share*
{EncryptedExtensions}
{Finished}

[Application Data*]

C→S:
{Finished}

[Application Data] <-----> [Application Data]

Режим PSK может использоваться в двух вариантах, определяемых расширением psk_key_exchange_modes: PSK и PSK-(EC)DHE. Последний дополняет процедуру формирования сеансовых ключей из PSK алгоритмом Диффи–Хеллмана (используется расширение key_share), что увеличивает как надежность итоговых ключей (обеспечивая прогрессирующую секретность, forwardsecrecy), так и сложность дополнительных криптографических вычислений.

Возобновление сеанса

Рассматриваемая версия протокола, в отличие от предыдущей версии, не выделяет отдельно режим возобновления сеанса. Вместо этого предлагается механизм согласования общего секрета PSK. Для этого используются так называемые удостоверения (tickets) и сообщение New Session Ticket, которое отправляется сервером клиенту в любое время после завершения сеанса рукопожатия и может использоваться в следующих соединениях. Сообщений New Session Ticket может быть несколько, каждое содержит одно удостоверение (ticket), которое с

одной стороны используется для формирования нового ключа PSK, с другой – является уникальным идентификатором этого ключа. Также при создании PSK используется соответствующий секрет из предыдущего сеанса рукопожатия, связывая ключевые материалы разных сеансов и обеспечивая дополнительную безопасность соединений.

Данные 0-RTT

На основе режима рукопожатия PSK реализован новый для протокола TLS способ отправки так называемых «ранних данных» (earlydata). При наличии общего ключа PSK, TLS 1.3 позволяет клиенту отправить некоторые данные во время первого рейса. Клиент использует PSK одновременно для аутентификации сервера и шифрования ранних данных.

C→S: ClientHello
+ early_data
+ key_share*
+ psk_key_exchange_modes
+ pre_shared_key
(Application Data*)

C←S: ServerHello
+ pre_shared_key
+ key_share*
{EncryptedExtensions}
+ early_data*
{Finished}
[Application Data*]

C→S:
(EndOfEarlyData)
{Finished}

[Application Data] <-----> [Application Data]

Таким образом, данные 0-RTT просто добавляются к рукопожатию 1-RTT в первом рейсе в сообщении Client Hello.

Как отмечено в спецификации, свойства безопасности для данных 0-RTT слабее, чем свойства безопасности для других данных TLS:

– Эти данные не обладают свойством прогрессирующей секретности, поскольку зашифрованы ключами, полученными с использованием только предлагаемого ключа PSK.

– Нет защиты от повторного воспроизведения между разными соединениями, поскольку данные 0-RTT не зависят от нового Server Hello (использующего случайное значение, Random). При этом внутри одного соединения данные 0-RTT дублироваться не могут, поскольку данные 0-RTT и 1-RTT защищаются разными ключами.

Обновление сеансовых ключей

Для повышения безопасности передачи данных рекомендуется периодически обновлять ключевой материал. Во время обмена рукопожатия создается базовый ключ соединения (mastersecret), из которого получают первоначальные сеансовые ключи. Сообщение Key Update сообщает партнеру, что отправитель создал новые сеансовые ключи, и последующие сообщения зашифрованы новыми ключами.

Расширение Cookie

Если в сообщении Client Hello недостаточно данных для продолжения рукопожатия, сервер может отправить сообщение Hello Retry Request, требующее прислать исправленное Client Hello с указанным ключевым материалом. При этом сервер может включить в Hello Retry Request необязательное расширение cookie [2]. Клиент должен скопировать это расширение в своем повторном Client Hello.

Как указано в спецификации, cookie преследует две основные цели:

- Заставляет клиента продемонстрировать достижимость на сетевом адресе (обеспечивая защиту от DoS-атак). Это прежде всего полезно для транспорта, не ориентированного на соединения.
- Позволяет серверу не сохранять состояние сеанса. Поскольку первое сообщение Client Hello (точнее, его хеш) участвует в криптографических вычислениях, сервер должен где-то его хранить. Вместо этого

сервер может выгрузить эти данные клиенту, включив их в расширение cookie, а клиент вернет его в ответном сообщении.

Расширение Signaturealgorithmscert

Чтобы указать, какие алгоритмы подписи могут использоваться в цифровой подписи, TLS 1.3 предоставляет два расширения [2]. Расширение “signature_algorithms_cert” применяется к подписям в сертификатах, а расширение “signature_algorithms” – к подписям в сообщениях Certificate Verify. В предыдущей версии TLS 1.2 используется только расширение “signature_algorithms”. Расширение “signature_algorithms_cert” было добавлено в новой версии протокола, чтобы позволить реализациям, поддерживающим разные наборы алгоритмов для сертификатов и для самого TLS, явно просигнализировать о своих возможностях. Ключи сертификатов должны соответствовать алгоритмам подписи, с которыми они используются. Если расширение “signature_algorithms_cert” отсутствует, то предполагается, что реализации используют одну и ту же политику в обоих случаях, и расширение “signature_algorithms” применяется также к подписям в сертификатах (что соответствует поведению TLS 1.2).

Расширение Servername

Могут возникнуть ситуации, когда по одному сетевому адресу размещено несколько «виртуальных» серверов, для каждого из которых используется свой сертификат. Данное расширение позволяет клиенту указать имя сервера, с которым он связывается. Для сервера получение этого расширения носит рекомендательный характер. Сервер может использовать информацию, содержащуюся в расширении, для выбора нужного сертификата, учитывая также и другие настройки политики безопасности. В этом случае серверу следует включить это расширение в соответствующее ответное сообщение. Хотя данное расширение представлено в отдельной более ранней спецификации, его поддержка и использование являются обязательными для реализаций TLS 1.3 [3].

Расширение Max fragment length / Recordsize limit

Если клиент хочет использовать сообщения меньшего размера, чем предусмотрено спецификацией (2^{14} байт), то он может воспользоваться расширением "max_fragment_length" [3]. Однако у него есть ряд существенных недостатков:

Данное расширение предлагает к применению всего несколько фиксированных значений, при этом максимальное значение (2^{12} байт) сильно меньше предлагаемого спецификацией TLS (2^{14} байт).

Добавление новых значений не предусмотрено.

Сервер не может запросить более низкое значение, чем то, которое предложил клиент. Это серьезная проблема, если сервер более ограничен.

Расширение плохо подходит для случаев, когда возможности клиента и сервера неодинаковы, т. е. если, например, конечная точка хочет отправлять записи большего размера, чем те, которые она получает.

Чтобы обойти эти ограничения, было предложено другое расширение для TLS – "recordsize limit" [4]. Оно позволяет задать произвольное максимальное значение сообщений (не превышающее значение, определенное согласованной партнерами версией протокола) для каждого направления передачи данных.

Расширение Posthandshakeauth

Расширение "post_handshake_auth" добавлено в TLS 1.3 и используется для указания того, что клиент хочет выполнить аутентификацию после рукопожатия [2]. В этом случае сервер может запросить аутентификацию клиента в любой момент времени после завершения рукопожатия путем посылки сообщения CertificateRequest. Клиент должен ответить соответствующими сообщениями аутентификации.

C←S: [CertificateRequest]

C→S:

[Certificate]

[CertificateVerify*]

[Finished]

Если у клиента нет соответствующих сертификатов, то отправляется пустое сообщение Certificate, а сообщение Certificate Verify не используется.

Сервер может отправить несколько сообщений Certificate Request, как в разное время, так и последовательно (например, если требуется доступ к нескольким сервисам). При этом ответы могут приходить в произвольной последовательности (для разделения запросов используются соответствующие уникальные идентификаторы).

Такая функциональность протокола может также использоваться и для режима рукопожатия PSK, во время которого сертификаты не используются, но зато после завершения рукопожатия можно запросить сертификат клиента (если ранее клиент включил расширение “post_handshake_auth” в сообщение Client Hello).

4. ВЕРИФИКАЦИЯ ПРОТОКОЛА

В процессе тестирования сетевых протоколов решается несколько важных задач: проверяется функциональная совместимость различных реализаций, проверяются соответствие реализации требованиям спецификации и устойчивость реализации к нестандартным воздействиям.

В наших проектах используются наработанные нами методики по тестированию сетевых протоколов: автоматизированное тестирование на соответствие формальным спецификациям и методы мутации данных.

В текущих экспериментах используется модель протокола TLS версии 1.3, разработанная нами на основе спецификаций RFC и описывающая сложную схему функционирования протокола.

Для тестирования реализаций на соответствие формальным спецификациям используется технология UniTESK, предоставляющая средства автоматизации тестирования на основе использования конечных автоматов [5]. Состояния тестируемой системы задают состояния автомата, а тестовые воздействия – переходы этого автомата. При выполнении перехода заданное воздействие передается на тестируемую реализацию, после чего регистрируются реакции реализации и автоматически выносятся вердикт о соответствии наблюдаемого поведения спецификации. В UniTESK алгоритм обхода конечного автомата реализован как внутренний компонент и не зависит от протокола и тестируемой системы.

Мутационные методы тестирования используются для обнаружения нестандартного поведения тестируемой системы (завершение из-за фатальной

ошибки, «подвисание», ошибки доступа к памяти). Как правило, подобные ситуации не рассматриваются в спецификациях. В поток обмена протокола, создаваемый в соответствии со спецификацией, вносятся некоторые изменения: либо в сообщениях, сформированных на основе разработанной модели протокола, изменяются значения полей, либо изменяется порядок сообщений в потоке обмена. Модель протокола позволяет изменять данные на любом этапе обмена, что позволяет тестовому сценарию проходить через все значимые состояния протокола и в каждом таком состоянии проводить тестирование реализации в соответствии с заданной программой.

5. ТЕСТОВЫЙ СТЕНД

Для тестирования реализаций сервера протокола TLS используются два сетевых узла. На одном узле функционирует модельная реализация под управлением UniTESK, выполняются основной поток управления тестовыми сценариями, обход тестового автомата и верификация наблюдаемых реакций. На другом узле функционирует тестируемая реализация. Тестовые сообщения протокола, сформированные модельной реализацией, передаются тестируемой системе, после чего регистрируются реакции тестируемого узла.

В качестве реализаций сервера TLSv1.3 используются:

- реализация TLS в виртуальной машине Java, JDK-14 (Java Development Kit) [6],
- реализация TLS библиотеки openssl-1.1.1f [7],
- реализация TLS в Windows 11 Pro 21H2 22000-100 (IIS).

Эти реализации являются частью широко используемых библиотек с открытым исходным кодом, имеют развитую функциональность и обеспечивают хорошую журнализацию событий.

6. РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ

Протокол TLSv1.3 позволяет добавлять новую функциональность за счет использования механизма расширений. В предыдущей части проекта акцент был сделан на использовании базовой функциональности протокола с минимальным количеством обязательных расширений. В данной работе используются как базовые

вые, так и дополнительные (описанные выше) возможности протокола. На данный момент в рамках технологии UniTESK (с использованием инструмента JavaTesK [8]) получены следующие результаты:

– разработана модель основной функциональности протокола TLS версии 1.3,

– разработаны спецификации и медиаторы для протокола TLS 1.3,

– разработан набор тестов, покрывающий часть требований спецификации.

Найдены несколько отклонений реализаций от спецификации, касающиеся обработки расширения cookie. Напомним, что в ответ на сообщение Client Hello, в котором отсутствует необходимый ключевой материал для выбранных сервером алгоритмов, сервер может отправить сообщение Hello Retry Request, требующее предоставить необходимые данные. После этого клиент должен прислать исправленное Client Hello, которое полностью повторяет первое, но с требуемым ключевым материалом. При этом, если сервер включил в Hello Retry Request необязательное расширение cookie, клиент должен скопировать это расширение в свое второе Client Hello.

JDK-14:

- сообщение Change Cipher Plain принимается после Client Finished (такие сообщения разрешается принимать только до Client Finished),
- сообщение от клиента с пользовательскими данными (тип Application) между сообщениями Server Finished и Client Finished принимается и игнорируется, хотя должны быть сообщения об ошибке (сообщения с типом Application клиент может отправлять только после Client Finished, при этом для шифрования Client Finished и Application используются разные ключи),
- сервер допускает несколько циклов сообщений Client Hello – Server Hello Retry (сообщение Server Hello Retry используется, если Client Hello предлагает приемлемые алгоритмы, но не хватает каких-либо данных; спецификация TLS разрешает только одно сообщение Server Hello Retry для каждого соединения),

- после обмена Рукопожатия (Handshake) сервер игнорирует тип входящих сообщений Протокола записей (Recordprotocol). Данный тип не несет никакой нагрузки, поскольку все сообщения зашифрованы, но в целях маскировки должен устанавливаться в значение для пользовательских данных (Application),
- странное поведение сервера: лишние данные после сообщения Client Hello остаются в буфере входящих сообщений сервера и в последующем рассматриваются как начальные байты следующего сообщения. При этом Client Hello не зашифровано, а остальные входящие сообщения должны быть зашифрованы. Данное поведение можно рассматривать как потенциальную уязвимость,
- реализация использует расширение cookie в сообщении HelloRetryRequest. При этом реализация успешно принимает ответное ClientHello с другим набором криптографических алгоритмов (поле TLS Cipher Suite), если клиент не включил в него расширениеcookie (если cookie присутствует, то реализация требует точного повторения всех полей, как и положено по спецификации),
- сообщение Hello Retry Request содержит алгоритм, для которого требуется прислать ключевой материал. Если клиент включил во второе ClientHello расширение cookie, то он может отправить ключевой материал для другого алгоритма (отличного от того, что запросил сервер). Если сервер поддерживает этот новый алгоритм, то сообщение принимается.
- Таким образом, в ответ на Hello Retry Request клиент может поменять в новом Client Hello предлагаемые алгоритмы. Однако такое поведение сервера может быть следствием того, что сервер не сохраняет первоначальное состояние сеанса (спецификация допускает работу сервера без сохранения состояния).

openssl-1.1.1f:

- не отвечает на сообщение Key Update с флагом 1, хотя для входящих от клиента сообщений создает новые ключи (сообщение Key Update уведомляет получателя, что отправитель изменил ключи для своих исходя-

щих сообщений, и следующие его сообщения защищаются новыми ключами; флаг 1 указывает, что получатель должен сделать тоже самое: отправить ответное сообщение Key Update и создать новые ключи для своих исходящих сообщений),

- реализация не использует расширение cookie в сообщении Hello Retry Request (что не запрещается), но при этом игнорирует присутствие и содержание данного расширения в ответном сообщении Client Hello,
- реализация игнорирует присутствие и содержание данного расширения в начальном сообщении Client Hello (расширение cookie используется клиентом только в повторном Client Hello в ответ на получение HelloRetryRequest).

InternetInformationServices,Windows 11:

- Сервер отвечает на сообщение Client Hello с единственным расширением Supported Versions <3.4> (спецификация требует присутствия в ClientHello еще нескольких обязательных расширений помимо Supported Versions, таким образом, серверу следовало бы ответить ошибкой и разорвать соединение).
- Сервер принимает Client Hello с дубликатами расширений (спецификация допускает присутствие только по одному расширению каждого типа).
- Сервер принимает от клиента сообщения Application и Record Layer Undefined Type Message до получения Client Finished (т. е. до завершения рукопожатия, эти сообщения уже зашифрованы сеансовыми ключами). Сервер игнорирует эти сообщения (хотя они нарушают допустимый порядок сообщений рукопожатия) и продолжает сеанс вместо разрыва соединения. Record Layer Undefined Type Message – тестовый тип сообщения, в котором в заголовке Record Layer поле Type выставлено в произвольное не зарегистрированное значение; Application – сообщение с пользовательскими данными.

ЗАКЛЮЧЕНИЕ

В работе представлен опыт верификации реализаций сервера криптографического протокола TLS версии 1.3. Протокол TLS является одним из наиболее широко используемых протоколов безопасности, обеспечивающих шифрование и защиту целостности данных при взаимной аутентификации сторон. Его используют в качестве транспорта многие другие прикладные протоколы и приложения (почтовые клиенты, интернет-браузеры, приложения голосовой и видеосвязи и др.). Появление версии 1.3 протокола TLS сделало актуальным проведение исследований в области верификации и безопасности реализаций этой новой версии протокола.

В данном проекте мы развиваем оригинальный подход к верификации реализаций и разработке тестового набора, сочетающий в себе два направления тестирования: тестирование реализаций протокола на соответствие требованиям стандарта и методы мутационного тестирования, предназначенные для обнаружения неадекватного поведения тестируемой системы (завершение из-за фатальной ошибки, «подвисание», ошибки доступа к памяти) в ситуациях, не определенных спецификацией.

Разработка модели протокола и тестового набора проводилась с использованием инструмента JavaTESK и языка программирования Java. С помощью этого нового тестового набора было проведено тестирование трех реализаций сервера протокола TLSv1.3: в виртуальной машине Java, JDK-14 (Java Development Kit), в библиотеке openssl-1.1.1f и реализация TLS в Windows 11 Pro 21H2 22000-100 (IIS). Результатом этих экспериментов стало обнаружение целого ряда отклонений этих реализаций от требований спецификации.

Таким образом, используемый нами подход к тестированию, совмещающий методы тестирования на соответствие требованиям спецификации и методы мутации данных, а также проведенные эксперименты по тестированию реализации протокола TLSv1.3 показали, что предложенный в данном проекте метод верификации позволяет эффективно автоматизировать тестирование; тестовые наборы обладают более полным покрытием требований и выявляют большее число ошибок, чем тестовые наборы, разработанные вручную или не использующие средств мутационного тестирования. Заметим также, что используемый нами

подход доказал свою эффективность и в наших предыдущих проектах при тестировании сетевых протоколов, обеспечив обнаружение различных отклонений от спецификации и других ошибок [9–11].

Исследование выполняется при поддержке РФФИ, проект № 20-07-00493 «Верификация функций безопасности и оценка устойчивости к атакам реализации протокола TLS версии 1.3».

СПИСОК ЛИТЕРАТУРЫ

1. *Dierks T., Rescorla E.* The Transport Layer Security (TLS) Protocol Version 1.2. August 2008. IETF RFC 5246. URL: <https://tools.ietf.org/html/rfc5246>
2. *Rescorla E.* The Transport Layer Security (TLS) Protocol Version 1.3. August 2018. IETF RFC 8446. URL: <https://tools.ietf.org/html/rfc8446>
3. *Eastlake D.* Transport Layer Security (TLS) Extensions: Extension Definitions. January 2011. IETF RFC 6066. URL: <https://tools.ietf.org/html/rfc6066>
4. *Thomson M.* Record Size Limit Extension for TLS. August 2018. IETF RFC 8449. URL: <https://tools.ietf.org/html/rfc8449>
5. *Bourdonov I., Kossatchev A., Kuliain V., and Petrenko A.* UniTesK Test Suite Architecture // Proceedings of FME 2002. LNCS 2391. P. 77–88, Springer-Verlag, 2002.
6. Java Development Kit 14.0.1 GA. URL: <https://jdk.java.net/14/>
7. OpenSSL Project. URL: <https://www.openssl.org/>
8. JavaTESK. URL: <http://www.unitesk.ru/content/category/5/25/60/>
9. *Никешин А.В., Пакулин Н.В., Шнитман В.З.* Разработка тестового набора для верификации реализаций протокола безопасности TLS // Труды ИСП РАН. 2012. Т. 23. С. 387–404.
10. *Никешин А.В., Пакулин Н.В., Шнитман В.З.* Тестирование реализаций клиента протокола TLS // Труды ИСП РАН. 2015. Т. 27, вып. 2. С. 145–160.
11. *Никешин А.В., Шнитман В.З.* Тестирование соответствия реализаций протокола EAP и его методов спецификациям Интернета // Труды ИСП РАН. 2018. Т. 30, вып. 6. С. 89–104. URL: [https://doi.org/10.15514/ISPRAS-2018-30\(6\)-5](https://doi.org/10.15514/ISPRAS-2018-30(6)-5)

EXPERIENCE OF IMPLEMENTATION OF THE PROTOCOL TLS 1.3 VERIFICATION

A. V. Nikeshin¹ [0000-0001-5781-9736], **V. Z. Shnitman**² [0000-0002-1509-0972]

Ivannikov Institute for System Programming of the Russian Academy of Sciences,
Alexander Solzhenitsyn st., 25, Moscow, 109004

¹alexn@ispras.ru, ²vzs@ispras.ru

Abstract

This paper presents the experience of verifying server implementations of the TLS cryptographic protocol version 1.3. TLS is a widely used cryptographic protocol designed to create secure data transmission channels and provides the necessary functionality for this: confidentiality of the transmitted data, data integrity, and authentication of the parties. The new version 1.3 of the TLS protocol was introduced in August 2018 and has a number of significant differences compared to the previous version 1.2. A number of TLS developers have already included support for the latest version in their implementations. These circumstances make it relevant to do research in the field of verification and security of the new TLS protocol implementations. We used a new test suite for verifying implementations of the TLS 1.3 for compliance with Internet specifications, developed on the basis of the RFC8446, using UniTESK technology and mutation testing methods. The current work is part of the TLS 1.3 protocol verification project and covers some of the additional functionality and optional protocol extensions. To test implementations for compliance with formal specifications, UniTESK technology is used, which provides testing automation tools based on the use of finite state machines. The states of the system under test define the states of the state machine, and the test effects are the transitions of this machine. When performing a transition, the specified impact is passed to the implementation under test, after which the implementation's reactions are recorded and a verdict is automatically made on the compliance of the observed behavior with the specification. Mutational testing methods are used to detect non-standard behavior of the system under test by transmitting incorrect data. Some changes are made to the protocol exchange flow created in accordance with the specification: either the values of the message fields formed on the basis of the developed protocol model are changed, or the order of messages in the

exchange flow is changed. The protocol model allows one to make changes to the data flow at any stage of the network exchange, which allows the test scenario to pass through all the significant states of the protocol and in each such state to test the implementation in accordance with the specified program. So far, several implementations have been found to deviate from the specification. The presented approach has proven effective in several of our projects when testing network protocols, providing detection of various deviations from the specification and other errors.

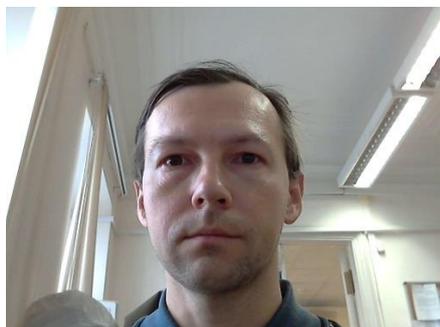
Keywords: *security, TLS, TLSv1.3, protocols, testing, verification, evaluate robustness, Internet, standards, formal specifications.*

REFERENCES

1. *Dierks T., Rescorla E.* The Transport Layer Security (TLS) Protocol Version 1.2. August 2008. IETF RFC 5246. URL: <https://tools.ietf.org/html/rfc5246>
2. *Rescorla E.* The Transport Layer Security (TLS) Protocol Version 1.3. August 2018. IETF RFC 8446. URL: <https://tools.ietf.org/html/rfc8446>
3. *Eastlake D.* Transport Layer Security (TLS) Extensions: Extension Definitions. January 2011. IETF RFC 6066. URL: <https://tools.ietf.org/html/rfc6066>
4. *Thomson M.* Record Size Limit Extension for TLS. August 2018. IETF RFC 8449. URL: <https://tools.ietf.org/html/rfc8449>
5. *Bourdonov I., Kossatchev A., Kuli Amin V., and Petrenko A.* UniTesK Test Suite Architecture // Proceedings of FME 2002. LNCS 2391. P. 77–88, Springer-Verlag, 2002.
6. Java Development Kit 14.0.1 GA. URL: <https://jdk.java.net/14/>
7. OpenSSL Project. URL: <https://www.openssl.org/>
8. JavaTESK. URL: <http://www.unitesk.ru/content/category/5/25/60/>
9. *Nikeshin A.V., Pakulin N.V., Shnitman V.Z.* Razrabotka testovogo nabora dlya verifikatsii realizatsiy protokola bezopasnosti TLS // Trudy ISP RAN /Proc. ISP RAS. 2012. Vol. 23. P. 387–404.
10. *Nikeshin A.V., Pakulin N.V., Shnitman V.Z.* TLS Clients Testing // Trudy ISP RAN /Proc. ISP RAS. 2015. Vol. 27, issue 2. P. 145–160.

11. *Nikeshin A.V., Shnitman V.Z.* Conformance testing of Extensible Authentication Protocol implementations // *Trudy ISP RAN/Proc. ISPRAS. 2018. Vol. 30, issue 6. P. 89–104 (in Russian).* URL: [https://doi.org/10.15514/ISPRAS-2018-30\(6\)-5](https://doi.org/10.15514/ISPRAS-2018-30(6)-5)

СВЕДЕНИЯ ОБ АВТОРАХ



НИКЕШИН Алексей Вячеславович – научный сотрудник Института системного программирования им. В.П. Иванникова РАН.

Aleksey Vyacheslavovich NIKESHIN – researcher, Ivannikov Institute for System Programming of the RAS.

Email: alexn@ispras.ru

ORCID: 0000-0001-5781-9736



ШНИТМАН Виктор Зиновьевич – д. т. н., с. н. с., заведующий отделом Института системного программирования им. В.П. Иванникова РАН.

Victor Zinovievich SHNITMAN – Doctor of Technical Sciences, Senior Research Officer, Head of Department, Ivannikov Institute for System Programming of the RAS.

Email: vzs@ispras.ru

ORCID: 0000-0002-1509-0972

Материал поступил в редакцию 22 октября 2021 года

УДК 519.178, 004.9

ИССЛЕДОВАНИЕ БИБЛИОГРАФИЧЕСКИХ ССЫЛОК В MATH-NET.RU С ИСПОЛЬЗОВАНИЕМ ГРАФА ЦИТИРОВАНИЯ ЖУРНАЛОВ

А. А. Печников¹, [0000-0002-0683-0019], Д. Е. Чебуков², [0000-0001-9738-8707]

¹Институт прикладных математических исследований – обособленное подразделение ФИЦ «Карельский научный центр Российской академии наук», г. Петрозаводск

² Математический институт им. В.А. Стеклова Российской академии наук, г. Москва

¹pechnikov@krc.karelia.ru, ²tche@mi-ras.ru

Аннотация

По данным портала Math-Net.Ru построен граф цитирования журналов, дугами в котором являются цитирования с 2002 по 2021 годы. Для повышения достоверности модели был выбран временной интервал цитирования с 2010 по 2021 годы, когда распределение цитируемых статей стабилизировалось на уровне 3500–4500 цитирований за год. Исследована структура старения ссылок и показано, что время их полужизни равно 8 годам. Поэтому дата издания цитируемых статей была ограничена 2002 годом. Для построенного графа цитирования получены основные свойства, такие как маленький диаметр и высокая плотность, свидетельствующие о высоком уровне научных коммуникаций в Math-Net.Ru. Показано отсутствие «эффекта Матфея» как ярко выраженного преимущества в цитировании состоявшихся журналов по отношению к менее известным. Адекватность графа цитирования журналов Math-Net.Ru как модели научных коммуникаций подтверждается сравнением ранжирования журналов в графе цитирования с их рейтингом SCIENCE INDEX в eLIBRARY.RU. Показана прямая умеренная связь между двумя ранжированиями. Сделан ряд содержательных выводов, следующих из анализа графа цитирования.

Ключевые слова: библиографическая ссылка, граф цитирования журналов, старение библиографических ссылок, индекс Матфея, математический портал Math-Net.Ru.

ВВЕДЕНИЕ

Одной из первых публикаций, посвященных анализу сетей научного цитирования, является статья де Солла Прайс [1], вышедшая в 1965 году и, несомненно, связанная с появлением электронных хранилищ научных публикаций. Основные исследуемые объекты с тех пор и до настоящего времени – сети цитирования и сети соавторства, построенные на основе данных, получаемых из различных электронных источников, включая Google Scholar, которые исследуются с использованием математических моделей и дальнейшей содержательной интерпретацией результатов.

Идея определения значимости научных журналов посредством измерения их цитируемости возникла гораздо раньше, еще в 1927 году [2], когда еще не использовалась термины типа «центральность в теории графов и анализе сетей». Заметим, что графом чаще называют математический объект (граф в строгом смысле как множество вершин и множество пар вершин, к которому применим аппарат теории графов), а сетью – те же структуры с содержательным наполнением (как сети цитирования журналов или сети соавторства). В контексте данной статьи эти два понятия практически идентичны. В работе [3] сказано, что «... изучение сетей цитирования как окна в науку представляет собой проверенную временем традицию», и дан хороший обзор публикаций на английском языке. Здесь используемые термины «граф» (graph) и «сеть» (network) также понимаются в похожем смысле.

В зарубежных исследованиях, посвященных анализу научного цитирования, основой для получения исходных данных являются крупные и всем известные поисковые интернет-платформы, такие как Web of Science (WoS) или Scopus, имеющие инструменты типа Journal Citation Reports, и (не менее крупные) специализированные ресурсы, такие как PubMed (URL: <https://pubmed.ncbi.nlm.nih.gov>). Русскоязычных журналов и публикаций на этих ресурсах немного, например, в WoS количество русскоязычных публикаций составляет менее 0.5% от общего количества.

В российских исследованиях сетей цитирования журналов большое внимание уделяется экономическому направлению [4–6]. Представляется, что во многом это связано с использованием базы данных RePEc (Research Papers in

Economics – URL: <http://repec.org>), значительная часть материалов которой находится в свободном доступе. Но и здесь русскоязычные журналы занимают незначительное место.

Небольшое число публикаций, связанных с исследованиями библиографических цитирований в России, возможно, связано с определёнными сложностями в получении необходимых сведений из баз данных Научной электронной библиотеки eLIBRARY.RU (URL: <https://elibrary.ru>).

В отличие от указанных случаев авторы имеют доступ к сведениям Общероссийского математического портала Math-Net.Ru (URL: <http://www.mathnet.ru>), что в определенной степени позволяет исследовать указанную проблематику в области математических наук в России (см., например, [7]).

Math-Net.Ru – известный веб-ресурс, содержащий богатую коллекцию полнотекстовых архивов ведущих российских математических журналов и информацию об их авторах. К середине марта 2021 года на портале зарегистрировано 133578 авторов, 295964 научных статьи из 143 журнала (периодических изданий).

По классификации eLIBRARY.RU более 85% журналов Math-Net.Ru относятся к тематике «математика», а остальные – к тематикам «информатика», «кибернетика» и «автоматика и вычислительная техника» (несколько журналов относятся сразу к нескольким из этих тематик). По данным eLIBRARY.RU за 2018 год из первых пятидесяти журналов с наибольшим Science Index по тематике «математика» 46 журналов входят Math-Net.Ru.

Информационная система Math-Net.Ru индексирует списки литературы и хранит их в базе данных в структурированном виде [8]. Списки литературы всех публикаций объединены в одну таблицу базы данных, в которой в отдельных колонках хранятся данные об авторе, названии, годе, томе, страницах цитируемой публикации. В несколько упрощенном виде запись о цитировании выглядит следующим образом:

<цитирующий журнал><цитирующая статья><цитируемый журнал><цитируемая статья>.

Данные о статьях, в свою очередь, имеют поля, позволяющие точно идентифицировать их: список авторов, название, год, страницы и т. д. Самая старая статья, проиндексированная в Math-Net.Ru, опубликована в 1866 году.

Каждой индивидуальной ссылке соответствует одна запись в таблице. Такой подход облегчает задачу автоматического проставления гиперссылок на библиометрические базы данных, решает задачу поиска обратных ссылок, а также позволяет автоматически экспортировать ссылки в разные форматы: PDF, XML, HTML. Среди гиперссылок с элементов списков литературы есть и ссылки на статьи, индексированные в базе публикаций Math-Net.Ru. Таким способом осуществляется связь между цитирующей и цитируемой статьями.

Основная задача данного исследования заключается в построении графа цитирования журналов Math-Net.Ru и его исследовании. Показана адекватность модели путем сравнения ранжирований для графа цитирования, eLibrary.Ru и Web of Science.

1. ОПРЕДЕЛЕНИЕ ВРЕМЕННОГО ИНТЕРВАЛА

В информационной системе Math-Net.Ru данные накапливаются в течение последних 15 лет, начиная с 2006 года. Самые ранние цитирующие статьи (то есть имеющие библиографические ссылки на публикации Math-Net.Ru), относятся к 1936 году, а самые ранние цитируемые статьи датируются 1866 годом.

Наполнение информационной системы по годам по различным причинам было достаточно неоднородным. Для достоверности построения модели цитирования журналов на первом этапе исследований из общего количества 143 журналов были отобраны те журналы, которые полностью индексируются в Math-Net.Ru, а соответствующие выпуски обработаны и открыты на сайте (для проверки в ручном режиме в случае необходимости).

Анализ распределения статей по годам показывает, что режим стабилизации прибавления новых публикаций наступил к 2010 году, начиная с которого в Math-Net.Ru стало добавляться от 3500 до 4500 статей ежегодно. Поэтому на первом этапе была сделана выборка цитирований из базы Math-Net.Ru с 2010 по 2021 год для статей, опубликованных в этих 143 журналах.

Общее количество цитирований за этот период равно 162069, и сделаны они с 41381 статьи на 61022 статей. Если цитирования суммировать по журналам, то оказывается, что за указанный период 106 цитирующих журналов ссылаются на 133 цитируемых журнала. Учитывая пересечения множеств цитирующих и цитируемых журналов, получаем множество из 135 журналов.

На рисунке 1 приведена гистограмма распределения статей, имеющих исходящие ссылки, по годам. Понятно, что 2021 год находится в стадии наполнения, поэтому количество статей всего 175. По остальным 11 годам имеем примерно одинаковую картину.

Очевидно, что возможно самоцитирование журналов [9]. К примеру, журнал «Алгебра и анализ» за 2010–2021 годы в своих статьях имеет 2517 библиографических ссылки на другие статьи, из которых в 542 случаях цитируются статьи из этого же журнала. Самоцитирование журналов – сложное явление, например, в работе [10] утверждается, что самоцитирование журналов требует их исключения из рассмотрения в индексах цитирования. Мы не придерживаемся столь категоричной позиции, считая, что самоцитирование журналов является особой формой публикационной научной коммуникации, и поэтому учитываем его в наших моделях.

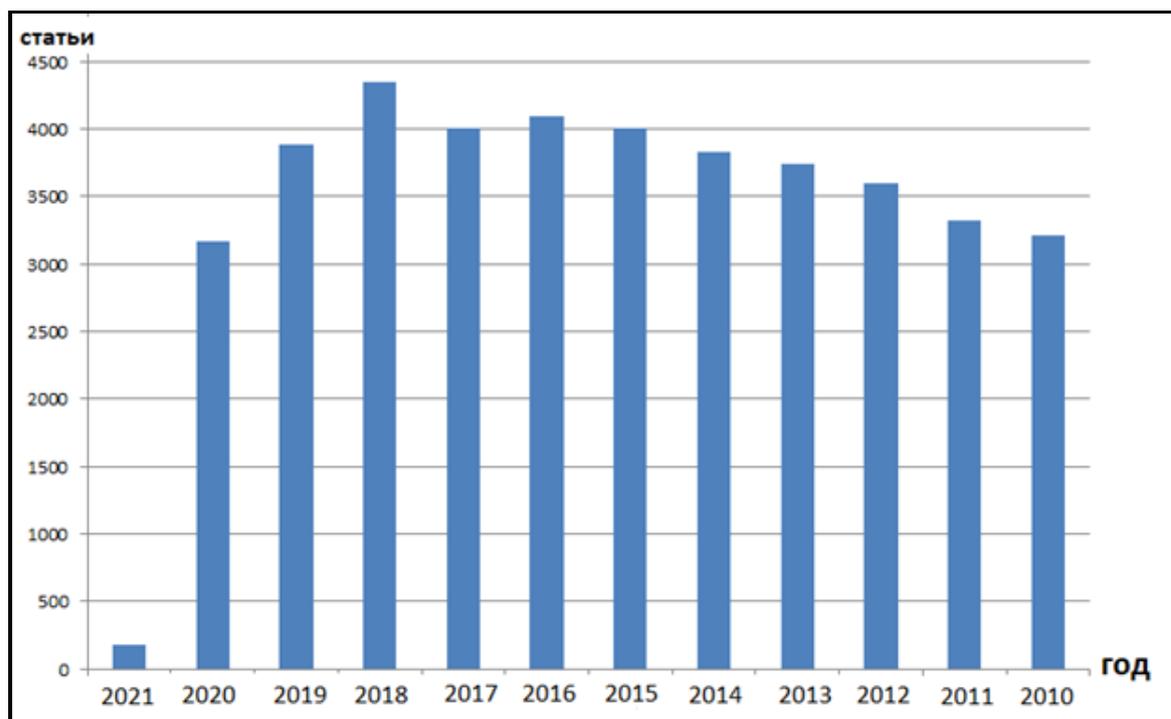


Рис. 1. Распределение цитирующих статей Math-Net.Ru в 2010–2021 годах

2. СТАРЕНИЕ ССЫЛОК

Для сформированного множества из 106 журналов, обязательно имеющих библиографические ссылки на другие журналы Math-Net.Ru за период 2010–2021 годов, была проанализирована возрастная структура ссылок.

Определим «возраст ссылки» как разницу между годом, когда опубликована статья, в которой приводится данная ссылка, и годом выпуска статьи, на которую ссылаются. «Время полужизни» цитируемости (или более точно «медиана хронологического распределения ссылок») определяется следующим образом: это такой момент времени, в который половина рассматриваемых ссылок относится к статьям, опубликованным позднее медианы, а другая половина – к статьям, более ранним, чем медиана.

На рисунке 2 показана интегральная возрастная структура исходящих ссылок данного множества журналов в виде графика, изображенного сплошной линией.

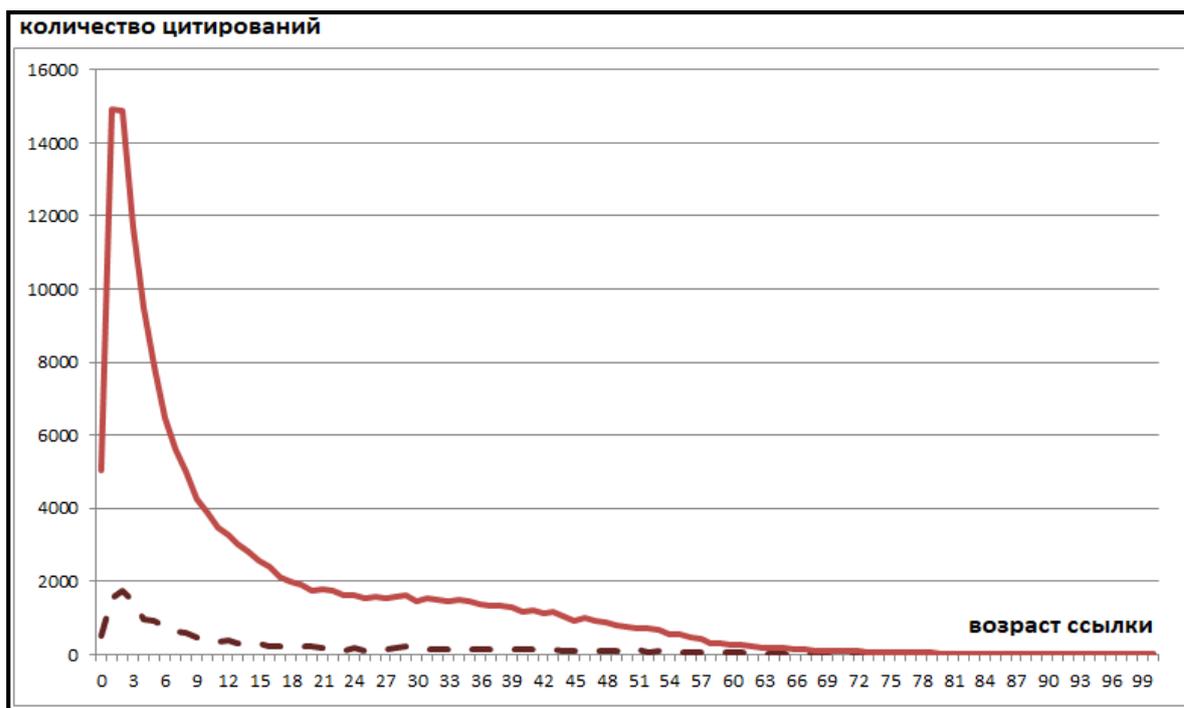


Рис. 2. Графики возрастной структуры исходящих ссылок

Например, значение в точке 0, равное 5060, вычисляется как суммарное количество всех ссылок, имеющих возраст, равный нулю. Возраст ссылки находится

в интервале от 0 до 152 лет. Максимальное количество цитирований (14924) приходится на статьи, опубликованные в предыдущем (по отношению к цитирующей публикации) году, то есть имеющие возраст цитирования, равный 1. Почти столько же цитирований приходится на возраст ссылки, равный 2, а затем происходит стремительный спад цитирований. Время полужизни ссылок для интегральной возрастной структуры равно 8 годам, то есть из 162 тысяч ссылок 81 тысяча имеет возраст не более 8 лет.

На этом же рисунке пунктирной линией изображен график возрастной структуры исходящих ссылок, сделанных в публикациях 2018 года. Значение в точке 0, равное 498, вычисляется как количество ссылок, сделанных в 2018 году на статьи, опубликованные в этом же году. Здесь максимум смещен на год правее, 1753 ссылки имеют возраст 2 года. Время полужизни ссылок 2018 года чуть меньше и равно 7,75 годам. Для сравнения, аналогичный график за 2010 год имеет практически такой же вид, за исключением того, что максимум достигается в точке 1, а время полужизни равно 9 годам.

3. СВОЙСТВА ГРАФА ЦИТИРОВАНИЯ

С учетом структуры старения и времени полужизни ссылок построим граф цитирования журналов Math-Net.Ru. Для этого будем использовать цитирования за 2010 по 2021 годы, ограничив годы издания цитируемых статей 2002 годом. Фактически мы используем найденный восьмилетний срок полужизни ссылок, и поэтому ссылки на статьи, опубликованные ранее 2002 года, не рассматриваются в нашей модели.

Указанное ограничение приводит к тому, что общее количество рассматриваемых журналов сократилось до 120, а количество цитирований – до 99 тысяч, причем почти 44 тысячи из них являются самоцитированиями. Уберем из множества журналов 13 журналов, имеющих только входящие или только исходящие ссылки, и сформированное множество журналов обозначим $J103$.

На множестве $J103$ построим граф цитирования журналов $G_{cit}(V, E, W)$, где:
 V – множество вершин (103 вершины, соответствующие журналам),
 E – множество дуг (3873 дуги, связывающие пары вершин i и j , если статьи журнала i имеют хотя бы одну ссылку на статьи журнала j),

W – множество весов дуг (вес $w(i,j)$ дуги $e(i,j)$ равен количеству ссылок, сделанных со всех статей журнала i на статьи журнала j).

Сумма всех весов W – это и есть количество всех цитирований журналов из $J103$, равное 97109, причем 43910 из них являются самоцитированиями. У большей части журналов количество самоцитирований больше, чем количество цитирований наиболее цитируемого журнала.

Укажем основные характеристики графа $Gcit(V,E,W)$. По построению очевидно, что он является сильно связным: из любой вершины в любую другую можно перейти по пути с конечным числом дуг. Диаметр графа (максимальное количество дуг в таком пути) равно 4. Плотность графа (отношение числа дуг к максимально возможному) достаточно велика и составляет 0,369.

Интересна структура модулярности графа $Gcit(V,E,W)$. Напомним, что графы с высокой модулярностью имеют сильные связи между вершинами внутри модулей, но слабые связи между узлами в различных модулях. Если рассматривать дуги без учета их весов, то модулярность практически равна нулю. Это означает, что дуги в графе распределены достаточно равномерно.

В случае учета весов дуг получаем несколько иную картину. Значение коэффициента модулярности в этом случае равно $Q=0,356$. Мы здесь используем определение меры модулярности Q из [11]. Значение Q лежит в интервале $[-1,1]$, и разбиение считается хорошим, если значение Q больше 0,7.

Полученное значение Q не показывает сильной модулярности графа $Gcit(V,E,W)$, но говорит о некоторых тенденциях. Граф $Gcit(V,E,W)$ разбивается на 5 модулей, каждый из которых может быть содержательно интерпретирован по направлениям исследований: фундаментальная математика, математическое моделирование, экспериментальная и теоретическая физика, дискретная математика и прикладная математика и компьютерные науки. Перечислим для примера журналы модуля «экспериментальная и теоретическая физика»: «Компьютерная оптика», «Международный научно-исследовательский журнал», «Письма в Журнал экспериментальной и теоретической физики», «Нелинейная динамика», «Квантовая электроника», “Regular and Chaotic Dynamics”, «Теплофизика высоких температур» и «Успехи физических наук».

Поскольку значимость научных журналов характеризуется посредством их ранжирования на основе показателей, построенных с использованием данных о

цитировании статей, граф цитирования также может быть использован для этих целей. Значимость вершин для ориентированного графа можно определить различными способами, и каждый из них требует содержательной интерпретации. Оценка Page Rank (PR) дает возможность сравнить относительную «важность» вершин графа по аналогии с важностью веб-страниц в Вебе [12]. Содержательная интерпретация значимости вершин по PR в графе $G(V,E,W)$ может быть следующей: если представить себе некоего «учёного-сёрфера», бесконечное число раз перемещающегося от одного журнала к другому по пристатейным ссылкам, то с наибольшей вероятностью он будет попадать на журнал с наибольшим значением PR .

Для графа $G(V,E,W)$ были вычислены значения PR для каждой вершины с учетом петель и весов. Значения PR , упорядоченные по убыванию значений для пяти первых и пяти последних вершин, приведены в таблице 1.

Таблица 1. Значения PR для графа $G(V,E,W)$

Вершина	Название журнала	PR
Mzm	Математические заметки	0,0607
Sm	Математический сборник	0,0516
Ufn	Успехи физических наук	0,0422
zvmmf	Журнал вычислительной математики и математической физики	0,0412
Smj	Сибирский математический журнал	0,0408
...
Co	Компьютерная оптика	0,0015
cgtm	Contributions to Game Theory and Management	0,0015
vtamu	Вестник российских университетов. Математика	0,0015
svfu	Математические заметки СВФУ	0,0015
thsp	Theory of Stochastic Processes	0,0015

4. ЭФФЕКТ МАТФЕЯ

Одним из показателей репутации журналов является так называемый «эффект Матфея», введённый Мертоном [13]. По евангельской цитате «... всякому имущему дастся и приумножится, а у неимущего отнимется и то, что имеет» (Мф. 25:29) в отношении научного цитирования эффект трактуется как преимущество в

цитировании состоявшихся ученых по отношению к их менее известным коллегам. В статье [14] эффект Матфея, как отсутствие преимущества в цитировании состоявшихся журналов по отношению к менее известным, проверялся для журналов Math-Net.Ru. Исследование проводилось для множества, состоящего из 166 журналов, проиндексированных в Math-Net.Ru за весь период времени, начиная от самых ранних проиндексированных публикаций и заканчивая 2019 годом. Оказалось, что в значительной степени наблюдается эффект, обратный эффекту Матфея, то есть «богатые богатеют ...», хотя «бедные не беднеют».

Поскольку в данной работе мы исследуем множество из 103 журналов, сформированное по специальным условиям, и временной интервал сокращен до периода после 2001 года, было интересно проверить, как эффект Матфея выполняется в этом случае.

Индекс Матфея для выявления отклонений в цитировании, связанных с государственной принадлежностью авторов, впервые был введен в [15]. По аналогии определим MI для журналов следующим образом:

$$MI_i = (obs_i - exp_i) / exp_i,$$

где obs_i – реальный прирост числа цитирований, а exp_i – ожидаемый прирост число цитирований статей, опубликованных в журнале i за период времени $[t_0, t_1]$. Здесь t_0 и t_1 – начальный и конечный моменты времени, определяющие интервал вычисления индекса Матфея. Как правило, t_0 и t_1 – это начальный и конечный годы интервала исследований.

Мы опишем вычисление MI_i применительно к нашему случаю. Для каждого журнала в базе данных определяется общее количество статей, опубликованное в нем и суммарное количество их цитирований от самого начала заполнения базы до года t_0 , что позволяет вычислить среднюю цитируемость статьи в журнале на год t_0 . Затем в базе данных определяется прирост количества статей, опубликованных в журнале после года t_0 до года t_1 , который умножается на среднюю цитируемость. Это и есть exp_i для i -го журнала. Реальный прирост числа цитирований obs_i за период времени $[t_0, t_1]$ определяется из базы данных.

Если MI_i больше 0, то делается вывод, что журнал i за период $[t_0, t_1]$ получает больше цитирований, чем можно предположить, и наоборот. Особенность этого

подхода, используемая в нашем случае, заключается в том, что нет необходимости определять состоявшиеся и менее известные журналы посредством некоего ранжирования, которое в свою очередь надо обосновывать.

Приведем в качестве примера один из характерных случаев, когда были взяты $t_0=2018$, а $t_1=2020$. Часть результатов приведена в таблице 2.

Таблица 2. Значения основных параметров и индекса Матфея

<i>Название журнала</i>	<i>cit18</i>	<i>p18</i>	<i>mid18</i>	<i>cit20</i>	<i>p20</i>	<i>exp_i</i>	<i>obs_i</i>	<i>MI_i</i>
Вестник Санкт-Петербургского университета. Серия 10. Прикладная математика. Информатика. Процессы управления	5478	95	57,7	6026	127	1845,2	548	- 0,70
Квантовая электроника	74358	1421	52,3	97870	1725	15907,7	23512	0,48
Письма в Журнал экспериментальной и теоретической физики	58457	1489	39,3	75566	1885	15546,6	17109	0,10
Теплофизика высоких температур	21212	1027	20,7	23417	1116	1838,2	2205	0,20
Regular and Chaotic Dynamics	5038	278	18,1	6308	335	1032,9	1270	0,23
Компьютерные исследования и моделирование	2513	143	17,6	3186	200	1001,6	673	- 0,33
Системы и средства информатики	3234	191	16,9	4114	244	897,4	880	- 0,02
Программные системы: теория и приложения	1554	94	16,5	1775	111	281,0	221	- 0,21
Препринты Института прикладной математики им. М.В. Келдыша РАН	8458	579	14,6	10345	731	2220,4	1887	- 0,15
Математическое моделирование	6213	490	12,7	8186	590	1267,9	1973	0,56

Здесь $cit18$, $p18$ и $cit20$, $p20$ – количество цитирований журнала и количество опубликованных в нем статей за 2002–2018 и 2002–2020 годы соответственно, а $mid18$ – среднее количество ссылок на статьи в 2018 году.

Журналы упорядочены по убыванию параметра $mid18$, то есть «Вестник Санкт-Петербургского университета» на 2018 год является самым «богатым» по этому показателю (по общему количеству цитирований таковым является «Квантовая электроника»). Как видно из таблицы, для первой десятки «богатых» по среднему количеству цитирований пять журналов имеют положительное значение индекса Матфея, а пять – отрицательное. Примерно такая же картина наблюдается и в целом с той разницей, что количество положительных значений индекса составляет около 80%. Поэтому можно сказать, что ярко выраженного преимущества в цитировании состоявшихся журналов по отношению к менее известным не наблюдается.

5. СРАВНЕНИЕ РАНЖИРОВАНИЯ ЖУРНАЛОВ В MATH-NET.RU, ELIBRARY.RU И WEB OF SCIENCE

Было проведено сравнение результатов ранжирования журналов Math-Net.Ru по показателю PR с результатами ранжирования в eLIBRARY.RU и Web of Science (WoS). Для этого в eLIBRARY.RU был взят такой показатель, как место журнала в рейтинге SCIENCE INDEX (SI) [16]. Для расчета SI журналы приписываются к одному из 10 тематических направлений. Для исследования было взято направление «Mathematics, computer and information sciences», куда попали 58 журналов из J103. В таблице 3 показаны десять первых и пять последних журналов, упорядоченных по убыванию в рейтинге SI, их значения PR и порядковые номера по PR . Неожиданным выглядит только высокое место по рейтингу SI у «Информатики и автоматизации».

Таблица 3. Значения *PR* для журналов, входящих в SCIENCE INDEX

<i>Название журнала</i>	<i>SI</i>	<i>PR</i>	<i>#PR</i>
Успехи математических наук	1	0,0384	5
Математический сборник	2	0,0516	2
Известия Российской академии наук. Серия математическая	3	0,0322	7
Математические заметки	4	0,0607	1
Информатика и автоматизация	5	0,0027	47
Алгебра и анализ	6	0,0226	9
Журнал вычислительной математики и математической физики	7	0,0412	3
Труды Математического института имени В.А. Стеклова	8	0,0340	6
Сибирский математический журнал	9	0,0408	4
Функциональный анализ и его приложения	10	0,0181	14
.....		
Математические заметки СВФУ	54	0,0015	58
Математическая физика и компьютерное моделирование	55	0,0016	56
Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления	56	0,0026	49
Вестник Тверского государственного университета. Серия: Прикладная математика	57	0,0017	53
Дагестанские электронные математические известия	58	0,0017	52

В качестве математического инструмента был использован ранговый коэффициент Спирмена. В основу метода положен принцип нумерации значений статистического ряда. Каждому элементу совокупности присваивается порядковый номер в ряду, который будет упорядочен по уровню признака (например, по убыванию). Таким образом, ряд значений признака ранжируется, а номер каждого элемента будет его рангом. Пусть n – количество наблюдаемых значений признака, x_i – ранг i -го элемента первого статистического ряда, y_i – ранг i -го элемента второго статистического ряда, а d_i – разность между рангами. Тогда ранговый коэффициент Спирмена равен

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}.$$

Связь принято считать сильной, если $|\rho| \geq 0.7$, средней силы для $0.5 < |\rho| \leq 0.69$.

В нашем случае в качестве x_i были взяты места журналов в рейтинге *SI*, а в качестве y_i – места в рейтинге *PR*. Полученное значение $\rho=0,66$ при критическом значении 0,01 свидетельствует об умеренной прямой связи между двумя ранжированиями.

Таблица 4. Значения *PR* для журналов, входящих в WoS

Название журнала	#WoS	IF WoS	#PR	PR
Успехи физических наук	1	2,821	20	0,0422
Письма в Журнал экспериментальной и теоретической физики	2	1,399	17	0,0391
Успехи математических наук	3	1,345	16	0,0384
Regular and Chaotic Dynamics	4	1,285	7	0,0143
Квантовая электроника	5	1,184	8	0,0162
Известия Российской академии наук. Серия математическая	6	1,13	14	0,0322
Теплофизика высоких температур	7	1,085	2	0,0055
Теоретическая и математическая физика	8	0,854	12	0,0256
Алгебра и анализ	9	0,8	11	0,0226
Математический сборник	10	0,8	21	0,0516
.....			
Moscow Mathematical Journal	18	0,544	5	0,0081
Функциональный анализ и его приложения	19	0,487	9	0,0181
Теория вероятностей и ее применения	20	0,485	6	0,0130
Труды Математического института имени В. А. Стеклова	21	0,467	15	0,0340
Журнал математической физики, анализа, геометрии	22	0,227	1	0,0028

В WoS входят 22 журнала из *J103*. Была сделана выборка из WoS значений их импакт-факторов [17] за 2019 год и проведено сравнение с выборкой их значений *PR*. В таблице 4 показаны десять первых и пять последних журналов, упорядоченных по убыванию в рейтинге WoS, их значения *PR* и порядковые номера по *PR*.

Значение рангового коэффициента Спирмена, полученное в этом случае, равно -0,29 и показывает отсутствие корреляции.

В первой десятке заметно выделяется половина физических журналов. Но и удаление из этого списка журналов физического направления оставляет корреляцию практически такой же.

ЗАКЛЮЧЕНИЕ

В информационной системе Math-Net.Ru хранятся публикации, датированные, начиная с 1886 года. Однако в целях повышения достоверности построенной модели для исследований был выбран временной интервал с 2010 по 2021 годы. Анализ структуры старения ссылок для этих публикаций показал, что время полужизни исходящих цитирований равно 8 годам. Поэтому для построения графа цитирования журналов Math-Net.Ru были взяты цитирования в статьях за 2010 по 2021 годы с ограничением даты издания цитируемых статей 2002 годом.

Для построенного графа цитирования получены основные свойства, к которым прежде всего можно отнести высокую плотность графа и маленький диаметр, равный 4, что свидетельствует о высоком уровне научных коммуникаций как журналов Math-Net.Ru, так и ученых, публикующихся в них. Значимость вершин графа определена с помощью показателя Page Rank. Отсутствие «эффекта Матфея» свидетельствует о высокой репутации исследуемого множества журналов.

Адекватность графа цитирования журналов Math-Net.Ru как модели научных коммуникаций подтверждается сравнением ранжирования журналов (как вершин в графе цитирования по PR) с их рейтингом SCIENCE INDEX в eLIBRARY.RU. Показана прямая умеренная связь между двумя ранжированиями.

В то же время аналогичная проверка с ранжированием в WoS показывает полное отсутствие корреляции. И это также косвенно подтверждает адекватность модели: большая часть журналов Math-Net.Ru не входит в WoS, и их взаимные цитирования никак не влияют на импакт-фактор WoS. На импакт-фактор, естественно, влияют цитирования, сделанные с публикаций, входящих в WoS, но они в большинстве случаев не входят в Math-Net.Ru и, по-видимому, не являются российскими.

В.В. Налимов определял наукометрию как «... количественные методы изучения развития науки как информационного процесса» [18]. Считается, что сегодня наукометрия применяется для оценки научной результативности и финансирования различных научных единиц (институтов, команд, индивидуумов). Поскольку эта работа относится к области наукометрии, сформулируем один из выводов, основанных на проведенных исследованиях.

Один из недавних официальных документов [19] «... определяет способ расчета значения качественного показателя, характеризующего публикационную результативность научных организаций,» и разработан «... в целях методического обеспечения формирования государственного задания».

По этой методике журнал «... без квартиля», входящий в WoS, имеет «коэффициент качества» примерно в 8 раз больший, чем журнал «... из списка ВАК» (Высшей аттестационной комиссии). По нашей модели получается, что высокий рейтинг, полученный в результате оценки российскими коллегами, практически ничего не значит по сравнению с оценками зарубежных коллег, если журнал не входит в WoS. И даже если он входит в WoS, его оценка формируется без учета мнения российского математического сообщества. Это очень напоминает известное изречение ««Поищемъ сами в собѣ князя, иже бы володѣль нами и рядиль по ряду, по праву» [20]. Получается, что хотя нами и не «володеют», но «рядят по ряду».

СПИСОК ЛИТЕРАТУРЫ

1. *D.J. de Solla Price. Networks of scientific paper // Science. 1965. No. 149(3683). P. 510–515.*
2. *Gross P.L.K., Gross E.M. College Libraries and Chemical Education // Science. 1927. V. 66, Iss. 1713. P. 385–389.*
3. *Kas M., Carley K.M., Carley L.R. Trends in science networks: understanding structures and statistics of scientific networks // Social Network Analysis and Mining. 2012. No. 2. P. 169–187.*
4. *Алескеров Ф.Т., Бадгаева Д.Н., Писляков В.В., Стерлигов И.А., Швыдун С.В. Значимость основных российских и международных экономических журналов: сетевой анализ // Журнал новой экономической ассоциации. 2016. № 2. С. 193–205.*

5. Бредихин С.В., Ляпунов В.М., Щербакова Н.Г. Структура сети цитирования научных журналов // Проблемы информатики. 2017. № 2(35). С. 38–51.
6. Бредихин С.В., Ляпунов В.М., Щербакова Н.Г. Спектральный анализ сети цитирования научных журналов // Проблемы информатики. 2018. № 2(39). С. 24–40.
7. Печников А.А., Чебуков Д.Е. Исследование графа научного сотрудничества на основе цитирования в Math-Net.Ru // Информатизация образования и науки. 2020. № 3(47). С. 98–106.
8. Chebukov D., Izaak A., Misyurina O., Pupyrev Yu., Zhizhchenko A. Math-Net.Ru as a digital archive of the Russian mathematical knowledge from the XIX century to today // Lecture Notes in Comput. Sci. 2013. Vol. 7961. P. 344–348.
9. Писляков В.В. Самоцитирование и его влияние на оценку научной деятельности: обзор литературы. Часть I // Препринт.
URL: <https://arxiv.org/ftp/arxiv/papers/2109/2109.09192.pdf>
10. Heneberg P. From Excessive Journal Self-Cites to Citation Stacking: Analysis of Journal Self-Citation Kinetics in Search for Journals, Which Boost Their Scientometric Indicators // PLoS ONE. 2016. 11(4): e0153730.
<https://doi.org/10.1371/journal.pone.0153730>
11. Newman M.E., Girvan M. Finding and evaluating community structure in networks // Physical Review E. 2004. Vol. 69(2). P 026113.
12. Brin S., Page L. The Anatomy of a Large-Scale Hypertextual Web Search Engine // Computer Networks and ISDN Systems. 1998. No. 30. P. 107–117.
13. Merton R.K. The Matthew Effect in Science // Science. 1968. Vol. 159. Iss. 3810. P. 56–63.
14. Печников А.А. О репутационном статусе журналов Math-Net.Ru // Cloud of Science. 2020. Т. 7. № 4. С. 734–748.
15. Bonitz M., Bruckner E., Scharnhorst A. The Matthew Index – Concentration Patterns and Matthew Core Journals // Scientometrics. 1999. Vol. 44, No. 3. P. 361–378.
16. Методика расчета интегрального показателя научного журнала в рейтинге Science Index. URL: https://elibrary.ru/help_title_rating.asp

17. The Clarivate Analytics Impact Factor.
URL: <https://clarivate.com/webofsciencegroup/essays/impact-factor>
18. *Налимов В.В., Мультченко З.М.* Наукометрия: Изучение развития науки как информационного процесса. М: Наука, 1969. 191 с.
19. Методика расчета качественного показателя государственного задания «Комплексный балл публикационной результативности» для научных организаций, подведомственных Министерству науки и высшего образования Российской Федерации, на 2020 год.
URL: https://old.minobrnauki.gov.ru/common/upload/library/2020/09/main/Methodika_novaya.pdf
20. Повесть временных лет. Подготовка текста, перевод и комментарии О.В. Творогова. URL: <http://lib.pushkinskiydom.ru/?tabid=4869>
-

RESEARCH OF BIBLIOGRAPHIC REFERENCES IN MATH-NET.RU USING THE JOURNAL CITATION GRAPH

A. A. Pechnikov¹ [0000-0002-0683-0019], D. E. Chebukov² [0000-0001-9738-8707]

¹ *Institute of Applied Mathematical Research of the Karelian Research Centre of the Russian Academy of Sciences, Petrozavodsk*

² *Steklov Mathematical Institute of RAS, Moscow*

¹pechnikov@krc.karelia.ru, ²tche@mi-ras.ru

Abstract

According to the portal Math-Net.Ru a graph of journal citation is constructed. To increase the reliability of the model, the citation time interval was chosen from 2010 to 2021, when the distribution of citation articles stabilized at the level of 3500-4500 citations per year. The structure of link aging is studied and it is shown that their half-life is equal to 8 years. Therefore, the publication date of the cited articles was limited to 2002. For the constructed citation graph, the main properties, such as a small diameter and a high density, are obtained, indicating a high level of scientific communication in the Math-Net.Ru. The absence of the “Matthew effect” as a pronounced advantage in quoting established journals in relation to less well-known ones is shown.

Adequacy of the journal citation graph Math-Net.Ru as a model of scientific communication confirmed by comparing the ranking of journals in the citation graph with their SCIENCE INDEX rating in eLIBRARY.RU. A direct moderate relationship between the two rankings is shown. A number of meaningful conclusions are drawn from the analysis of the citation graph.

Keywords: *bibliographic reference, journal citation networks, aging of bibliographic references, Matthew index, mathematical portal Math-Net.Ru.*

REFERENCES

1. *D.J. de Solla Price.* Networks of scientific paper // *Science.* 1965. No. 149(3683). P. 510–515.
2. *Gross P.L.K., Gross E.M.* College Libraries and Chemical Education // *Science.* 1927. V. 66, Iss. 1713. P. 385–389.
3. *Kas M., Carley K.M., Carley L.R.* Trends in science networks: understanding structures and statistics of scientific networks // *Social Network Analysis and Mining.* 2012. No. 2. P. 169–187.
4. *Aleskerov F.T., Badgaeva D.N., Pislyakov V.V., Sterligov I.A., Shvydun S.V.* Znachimost' osnovnyh rossiiskih i mejdunarudnyh ekonomicheskikh jurnalov: setevoi analiz // *Journal of the New Economic Association.* 2016. № 2. P. 193–205.
5. *Bredikhin S.V., Lyapunov V.M., Shcherbakova N.G.* The structure of citation network of scientific journals // *Problems of Informatics.* 2017. №2 (35). P. 38–52.
6. *Bredikhin S.V., Lyapunov V.M., Shcherbakova N.G.* Spectral'nyi analiz seti citirovanija nauchnyh jurnalov // *Problems of Informatics.* 2018. № 2(39). P. 24–40.
7. *Pechnikov A., Chebukov D.* Issledovanie grafa nauchnogo sotrudnichestva na osnove citirovanija v Math-Net.Ru // *Informatizacija obrazovanija I nauki.* № 3(47). 2020. S. 98–106.
8. *Chebukov D., Izaak A., Misyurina O., Pupyrev Yu., Zhizhchenko A.* Math-Net.Ru as a digital archive of the Russian mathematical knowledge from the XIX century to today // *Lecture Notes in Comput. Sci.* 2013. Vol. 7961. P. 344–348.
9. *Pislyakov V.* Self-Citation and its Impact on Research Evaluation: Literature Review. Part I // Preprint.
URL: <https://arxiv.org/ftp/arxiv/papers/2109/2109.09192.pdf>.

10. *Heneberg P.* From Excessive Journal Self-Cites to Citation Stacking: Analysis of Journal Self-Citation Kinetics in Search for Journals, Which Boost Their Scientometric Indicators // PLoS ONE. 2016. 11(4): e0153730.
<https://doi.org/10.1371/journal.pone.0153730>
11. *Newman M.E., Girvan M.* Finding and evaluating community structure in networks // Physical Review E. 2004. Vol. 69(2). P 026113.
12. *Brin S., Page L.* The Anatomy of a Large-Scale Hypertextual Web Search Engine // Computer Networks and ISDN Systems. 1998. No. 30. P. 107–117.
13. *Merton R.K.* The Matthew Effect in Science // Science. 1968. Vol. 159. Iss. 3810. P. 56–63.
14. *Pechnikov A.A.* On the reputation status of Math-Net.Ru journals // Cloud of Science. 2020. Vol. 7. No. 4. P. 734–748.
15. *Bonitz M., Bruckner E., Scharnhorst A.* The Matthew Index – Concentration Patterns and Matthew Core Journals // Scientometrics. 1999. Vol. 44, No. 3. P. 361–378.
16. Metodika rascheta integral'nogo pokazatelja nauchnogo jurnala v reitinge Science Index. From URL: https://elibrary.ru/help_title_rating.asp (in Russian).
17. The Clarivate Analytics Impact Factor.
URL: <https://clarivate.com/webofsciencegroup/essays/impact-factor>
18. *Nalimov V., Mul'chenko Z.* Naukometrija: Izuchenie razvitija nauki kak informacionnogo processa. M.: Nauka, 1969. 191 s.
19. Metodika rascheta kachestvennogo pokazatelja gosudarstvennogo zadanija "Kompleksnyi ball publikacionnoi rezul'tativnosti" dlja nauchnyh organizacii, podvedomstvennyh Ministerstvu nauki i vys'shego obrazovanija Rossiiskoi Federacii, na 2020 god.
URL: https://old.minobrnauki.gov.ru/common/upload/library/2020/09/main/Metodika_novaya.pdf
20. Povest' vremennyh let. Podgotovka teksta, perevod i kommentarii O.V. Tvorogova. URL: <http://lib.pushkinskijdom.ru/?tabid=4869>

СВЕДЕНИЯ ОБ АВТОРАХ



ПЕЧНИКОВ Андрей Анатольевич – главный научный сотрудник Института прикладных математических исследований – обособленного подразделения ФИЦ «Карельский научный центр Российской академии наук», д. т. н. Сфера научных интересов – математическое моделирование, дискретная оптимизация, вебометрика, наукометрия.

Andrey Anatolievich PECHNIKOV – Chief Research Associate, Institute of Applied Mathematical Research of the Karelian Research Centre of the Russian Academy of Sciences, Doctor (DSc) of Technics. Research interests include mathematical modeling, discrete optimization, webometrics, scientometrics.

email: pechnikov@krc.karelia.ru,

ORCID: 0000-0002-0683-0019



ЧЕБУКОВ Дмитрий Евгеньевич – зав. информационно-издательским сектором Математического института им. В.А. Стеклова Российской академии наук, к. х. н. Сфера научных интересов – библиометрия, наукометрия, электронные библиотеки.

Dmitry Evgen'evich CHEBUKOV – Head of Information and Publishing Sector, Steklov Mathematical Institute of Russian Academy of Sciences, Candidate Chem. Sci. Research interests include bibliometrics, scientometrics, digital libraries.

email: tche@mi-ras.ru,

ORCID: 0000-0001-9738-8707

Материал поступил в редакцию 12 октября 2021 года

ЗАМЕТКИ О ПЕРВЫХ ЭЛЕКТРОННЫХ БИБЛИОТЕКАХ

Ю. Е. Поляк^[0000-0001-8411-335X]

Федеральное государственное бюджетное учреждение науки Центральный экономико-математический институт Российской академии наук, 117418
Москва, Нахимовский пр., д. 47

polak@semi.rssi.ru

Аннотация

Электронные библиотеки в XXI веке стали важнейшими источниками информации во всех отраслях науки и техники, магистральным направлением развития мирового библиотечного дела, всей информационной деятельности. Первой электронной библиотекой стал Проект Гутенберг, основанный Майклом Хартом ровно 50 лет назад. Всемирно известным он стал лишь в 1990-е годы, после появления интернета и первых браузеров. Харт мечтал о 10000 оцифрованных книг и, по его признанию, достиг всех целей, которые ставил перед собой. В России одной из первых и наиболее значительной стала библиотека Максима Мошкова (1994). В данной статье прослеживаются этапы создания этих библиотек и биографии их создателей; анализируются трудности в их работе. Отдельное внимание уделено вопросам законодательства о продолжительности действия авторских прав.

Ключевые слова: электронные библиотеки, Проект Гутенберг, Майкл Харт, библиотека Мошкова, авторские права.

ВВЕДЕНИЕ

Электронные библиотеки – неотъемлемая часть современной жизни. Однако первая такая библиотека появилась в некотором смысле случайно.

4 июля 1971 года студент университета штата Иллинойс Майкл Харт получил машинное время на суперкомпьютере Хегох Sigma V в университетской лаборатории исследования материалов на сумму, эквивалентную \$100 миллионам (по другим данным – 100 тысячам). Он оказался в нужном месте в нужное время: операторы были друзьями Майкла и его брата, а на новом мэйнфрейме было много

свободного времени, и операторов поощряли заниматься чем угодно, чтобы они повышали квалификацию.

Харт понял, что стал обладателем огромного богатства, и задумался, как использовать неограниченное количество компьютерного времени для чего-то более полезного, чем банальные вычисления. Раздумья продолжались, по его словам, час и 47 минут (напомним, Гагарину в 1961 году потребовалось на полёт чуть больше – 108 минут). Говорит сам Харт (интервью 2009 г.): «4 июля, ещё будучи первокурсником Иллинойского университета, я после праздничного фейерверка решил переночевать за мэйнфреймом Xerox Sigma V в исследовательском центре вместо того, чтобы идти пешком домой в летнюю жару, а через несколько часов возвращаться на занятия. По пути я зашёл в магазин сети IGA купить на ночь что-нибудь поесть. Вместе с продуктами мне положили копию Декларации независимости США на искусственном пергаменте. В ту ночь я получил свой первый компьютерный аккаунт на огромную сумму и решил, что должен сделать что-то чрезвычайно ценное. Это был такой напряжённый мыслительный процесс для первокурсника, что моей первой мыслью было съесть что-нибудь, чтобы набраться сил. Вместе с продуктами выпал пергамент, и меня озарила идея, буквально свет загорелся над моей головой, как в мультфильмах. Я понял, каким будет будущее компьютеров и интернета. В результате я сел и напечатал этот исторический документ, и в ночь 4 июля 1971 года появился Проект Гутенберг. Остальное уже история»¹. В ту ночь Майкл Харт решил, что «наибольшей ценностью, создаваемой компьютерами, будут не вычисления, а хранение и поиск того, что находится в наших библиотеках».

Мэйнфрейм того времени разительно отличался от привычных современных компьютеров: у него не было ни монитора, ни клавиатуры. Харту пришлось набирать текст на другом устройстве – телетайпе, преобразовывавшем буквы в компьютерный код путём пробивания отверстий в бумажной ленте, которая вставлялась в ЭВМ. Текст Декларации независимости был набран ПРОПИСНЫМИ БУКВАМИ, нижний регистр эта техника «не понимала». Документ получился довольно большим – опять-таки по меркам 1971 года – целых 5 килобайт.

¹ <https://www.gutenberg.org/files/36616/36616-pdf.pdf>

У иллинойского компьютера Xerox Sigma V была важная особенность: он был одним из 15 узлов компьютерной сети ARPANet, которая впоследствии переросла в знакомый всем интернет. Харт собирался разослать Декларацию независимости каждому пользователю сети, но его предупредили, что отправка такого большого текста в сотню адресов может привести к сбою работы всей сети. По этой причине Харт ограничился рассылкой информации, позволявшей узнать, где в системе хранилась электронная версия Декларации независимости (или файл «e-book», как он его назвал). «Я разместил сообщение на том, что позже стало comp.gen, группой новостей Usenet, и хранил файлы на магнитной ленте. Люди отправляли электронное письмо, и кто-то в компьютерной лаборатории, обычно мой лучший друг, загружал ленту. Затем мы сообщали им, где находится файл и как его скопировать».

Вспоминает один из пользователей. «В 1971 году, когда г-н Харт официально начал проект Гутенберг, или даже в начале 1990-х годов, когда проект начал достигать критической массы, электронные книги были не просто необычными – они были неслыханными. Помню волшебство чтения Декларации независимости через telnet-соединение, задолго до того, как интернет сделал такие вещи тривиальными. Тогда это было так же здорово, как путешествие по миру через серверы gopher несколькими годами позже. До ePub, электронных книг, до Википедии, даже до интернета, г-н Харт увидел, что цифровые коммуникации могут изменить гуманитарные науки и весь мир»². Об ftp-архивах, протоколах gopher, telnet и других «довебовских» сетевых технологиях см. «Big Dummy's Guide to the Internet»³, 1993 (русский перевод⁴ 1994 г.), а также [3]. Кстати, один абзац этого справочного руководства посвящён Проекту Гутенберг⁵.

Из примерно сотни участников тогдашней Сети шесть человек откликнулись на призыв Харта и скопировали Декларацию независимости на свои компьютеры. Она, таким образом, стала первой электронной книгой, доступной по электронным сетям. За ней последовали билль о правах, речи Д. Кеннеди и А. Линкольна,

² <https://www.chronicle.com/article/michael-hart-1947-2011-defined-the-landscape-of-digital-publishing>

³ <https://science.ksc.nasa.gov/facts/internet/bdgtti-1.01.html>

⁴ http://citforum.ru/internet/inet_tut

⁵ http://citforum.ru/internet/inet_tut/7_8.shtml

конституция США, Библия, Приключения Алисы в стране чудес ... Электронной книгой №100 стало полное собрание сочинений Шекспира⁶ (размещено в январе 1994 г.; последняя правка на момент подготовки этого текста относится к августу 2021 г.). Всё это Харт оцифровывал собственноручно, что требовало огромных усилий и времени. Но им руководила амбициозная цель – сделать к 2001 году доступными через сеть 10000 книг и документов. Так было положено начало Проекту Гутенберг, крупнейшей для своего времени коллекции бесплатных электронных книг. Выбрав это название, Майкл Харт указал масштаб своего проекта, революционного для распространения литературы, как и изобретение подвижного типа печати в 1450-х годах.

ГУТЕНБЕРГ



Немецкий первопечатник, первый типограф Европы Иоганн Генсфляйш цур Ладен цум Гутенберг (Johannes Gensfleisch zur Laden zum Gutenberg) родился в Майнце около 1400 года. Записей о его крещении не сохранилось; точно известна лишь дата смерти – 3 февраля 1468 г. В молодости он обучался ювелирному делу, занимался шлифовкой драгоценных камней и изготовлением зеркал, параллельно работая над изобретением, которое держал в секрете. В 1439 году он впервые в Европе стал использовать набор подвижных литер, т. е. выпуклых зеркальных отображений отдельных букв, составлявшихся в нужном порядке. В документах того времени упоминаются покупка им свинца и строительство винтового пресса – прототипа печатного станка. После 1448 года Иоганн Гутенберг начинает эксперименты с печатью. Около 1450 года он изобрёл и самостоятельно разработал европейскую систему подвижных литер, дополнив её такими новшествами, как изготовление шрифта при помощи пуансонов и матриц. Гутенберг был первым, кто начал делать литеры из сплава свинца, олова и сурьмы – те же компоненты использовались в течение последующих 400 лет. Вначале из-под его печатного пресса выходят недорогие издания на одном листе –

⁶ <https://www.gutenberg.org/files/100/100-0.txt>

индальгенции, календари, затем латинская грамматика. А в 1455 году было напечатано 180 экземпляров двухтомной Библии – замечательного образца полиграфического искусства. В этом легко позволяет убедиться иллюстрация: экземпляр, хранящийся в Российской государственной библиотеке, оцифрован и общедоступен⁷.



Высокое качество и относительно низкая цена книг Гутенберга имели важные социальные последствия. С появлением механической печати начались резкий рост грамотности и обмен информацией. Была разрушена монополия элиты на образование и обучение, уменьшилось влияние политических и религиозных властей, получил поддержку нарождающийся средний класс. Эпоха массовой коммуникации навсегда изменила структуру общества.

Строго говоря, Гутенберг не был первооткрывателем такого способа печати.

Впервые подвижной шрифт изобрёл китаец Би Шэн около 1040 года. Подвижные литеры из металла впервые были применены в 1234 году в Корее. Ещё до рождения Гутенберга, в 1377 году в Корее появилась Чикчи (Антология учения великих монахов об обретении духа Будды) – старейшая из существующих книг, напечатанных при помощи металлических подвижных литер. Однако распространение этих первых систем было ограниченным: они были дорогими и требовали больших трудозатрат. Китайским печатникам приходилось иметь дело с множеством керамических табличек, поскольку иероглифическая система письма содержит тысячи знаков⁸.

В алфавитных системах письма набор посредством подвижных литер оказался более быстрым и надёжным, чем ксилография. Металлические литеры были прочнее, а текст получался единообразным, что привело к появлению типографии и шрифтового искусства. Появилась практическая система, позволявшая

⁷ https://viewer.rusneb.ru/ru/000199_000050_000840779;

https://viewer.rusneb.ru/ru/000199_000050_000840780

⁸ https://ru.wikipedia.org/wiki/Пучной_набор

массово выпускать печатные книги и экономически выгодная и для типографов, и для читателей. Гутенберговская печать подвижными литерами оказала огромное влияние не только на европейскую культуру, но и на всемирную историю. Печатный станок ускорил наступление научной революции, эпох Ренессанса, Реформации, Просвещения и заложил материальную основу для современной экономики, основанной на знаниях и распространении обучения в массах.

ПЕРВОПЕЧАТНИК ИНТЕРНЕТА. ЖИЗНЬ МАЙКЛА ХАРТА

Майкл Харт (*Michael Stern Hart*) родился в Такоме, штат Вашингтон, 8 марта 1947 года. Его мать, имевшая три диплома, во время Второй мировой войны работала криптоаналитиком, а впоследствии была бизнес-менеджером в «самом шикарном женском магазине» города. Отец работал бухгалтером. Когда Харту было одиннадцать, семья переехала в город Урбана, административный центр округа Шампейн штата Иллинойс (американцы произносят Эрбана и Чемпейн; Urbana [z:r' bænə], Champaign ['tʃæmpreɪn]). Родители Майкла стали профессорами колледжей шекспировских исследований и математического образования местного университета (отец и мать соответственно)⁹. Учась в школе, Майкл был «скаут-орлом» (Eagle Scout – высший ранг американских бойскаутов)¹⁰. Во время вьетнамской войны он служил в армии в Корее. Затем он учился в Университете штата Иллинойс, который окончил за два года со степенью в области человеко-машинных интерфейсов, потом поступил в аспирантуру, но не окончил её. Некоторое время Харт был уличным музыкантом¹¹, но в итоге предпочёл заниматься техническими науками.

⁹ http://lib-rarium.blogspot.com/2012/12/1_19.html

¹⁰ [https://en.wikipedia.org/wiki/Eagle_Scout_\(Boy_Scouts_of_America\)](https://en.wikipedia.org/wiki/Eagle_Scout_(Boy_Scouts_of_America))

¹¹ https://www.wikiwand.com/en/Michael_S._Hart

Он всю жизнь что-то чинил и мастерил. В детстве разбирал телевизоры и радиоприёмники, чтобы посмотреть, как они устроены, после чего собирал их обратно. Позднее экспериментировал с компьютерными технологиями, начиная с первых персоналок Apple и Atari, операционных систем CP/M, MS-DOS и других. Эти навыки впоследствии ему очень пригодились.



С того июльского вечера, когда стартовал Проект Гутенберг, на протяжении следующих 40 лет вся дальнейшая жизнь Майкла Харта, энтузиаста и бесребренника, была посвящена созданию открытой библиотеки и распространению литературы в электронном виде. Большую часть времени он проводил в подвальном этаже своего дома, заваленном книгами.

Харт жертвовал личной роскошью ради борьбы за грамотность и сохранение ресурсов общественного достояния. Доходов это не приносило, он прожил жизнь практически в бедности. Но это не мешало ему сохранять хорошее настроение, оптимизм и уважение к другим. Вот его слова из интервью журналисту Р. Пойндеру 2006 года¹².

«— Получаете ли Вы зарплату от фонда литературного архива Проекта Гутенберг, который был основан в 2000 году?

— Нет. Мы не привлекаем для этого достаточно средств.

— А когда закончатся сбережения?

— Я не собираюсь позволить им закончиться. Я либо выйду из строя, либо к тому времени получу больше денег. Для многих это прозвучит странно, но деньги никогда не имели для меня большого значения, я никогда не тратил их, когда получал. Поскольку я трачу немного и всё откладываю, деньги вообще не проблема. Большинство людей тратят деньги на вещи, которые меня просто не волнуют. Я замечаю это только тогда, когда думаю о покупке нового дома.

— Значит, Вы ведёте несколько спартанский образ жизни?

¹² https://ia802307.us.archive.org/23/items/The_Basement_Interviews/Michael_Hart_Interview.pdf

– Для меня сложно потратить на обед \$10; как правило, эта сумма меньше \$5. У меня нет кабельного телевидения и сотового телефона. Обычно я передвигаюсь на велосипеде и покупаю машину, только если мне предстоит долгая поездка; тогда я покупаю дешёвый автомобиль, и он мне служит ещё пару лет. Вещи я покупаю на гаражных распродажах. В нашем университетском городке иногда бывает более 200 таких распродаж в один день. Гаражные распродажи также дают отличную возможность познакомиться с людьми со всего города, поискать их вещи и поговорить обо всём».

Харт зарабатывал на жизнь случайными заработками; профессиональные связи помогали ему привлекать дополнительные инвестиции, гранты и пожертвования. Расходы его были сведены к минимуму: он собирал компьютеры, стереосистемы и другое оборудование из выброшенных компонентов, использовал домашние средства вместо обращения к врачам, самостоятельно ремонтировал собственный дом и машину, питался бобовыми консервами. По протекции капеллана при Иллинойском университете Дэвида Тёрнера он получил должность адъюнкт-профессора Бенедиктинского колледжа в Иллинойсе со стипендией \$1000 в месяц. Эти деньги Харт вкладывал в свой проект.

Обратимся к интервью 2002 года (с Хартом беседовал С. Вакнин, назвавший этот текст «Второй Гутенберг», а самого Харта – провидцем, опередившим своё время, и кладезем блестящих идей)¹³.

«– В каком смысле этот проект сравним с революцией Гутенберга?

– Когда я выбирал название, главным фактором было то, что публикация электронных книг меняет картину грамотности и образования так же, как это сделал пресс Гутенберга, который снизил цену книг в 400 раз. Эта цена была сопоставима со стоимостью семейной фермы, а потом упала до того, что книги появились на воскресных рынках в небольших деревнях. Когда у нас была только дюжина электронных книг в интернете, цена размещения каждой из них на компьютере составляла примерно 1/400 цены книги в мягкой обложке. А когда 100-гигабайтные диски подешевели до \$100, это стало «слишком дешево для измерения». Те, кто меряет всё деньгами, невероятно расстроены Проектом Гутенберг. Ещё одно

¹³ <https://samvak.tripod.com/busiweb29.html>

сходство нашего проекта с революцией Гутенберга заключается в том, что законы об авторских правах были созданы, чтобы остановить и то, и другое. Проект Гутенберг – это первый пример смены парадигмы от ограниченного к неограниченному распространению, рекламируемому сейчас как «век информации». Однако это уже четвёртый информационный век. Каждая такая фаза была подавлена, использование новых технологий для копирования текстов объявлялось незаконным».

Как видим, Харт не скрывает существование недоброжелателей, чьи коммерческие интересы проект ущемил. По его мнению, «протекционизм препятствует распространению грамотности, лишает массы столь необходимых знаний, дискриминирует бедных и, в конечном счете, подрывает демократию». Это диаметрально противоположно цели Проекта Гутенберг – стимулировать создание и распространение информации для снижения информационной неграмотности среди пользователей. Книги проекта также могут помочь преодолеть «цифровой разрыв». Их можно читать даже в карманном компьютере на солнечных батареях ценой в несколько долларов, что актуально для отдалённых регионов и развивающихся стран. В цитированном интервью Харт говорил: «Из Африки мне сообщают, что там люди читают и слушают электронные книги на своих мобильных телефонах. В это трудно поверить, но, когда вы находитесь на равнинах Серенгети, вы всё равно получаете услуги сотовой связи. Существуют миллиарды сотовых телефонов; миллиарды ноутбуков и миллиарды настольных компьютеров. Но пока нет даже миллионов читателей электронных книг». И далее: «Я аполитичен, вплоть до антиполитичности. Я вообще не верю в современный политический процесс. Я не устанавливаю политику, я просто пытаюсь подавать пример. Электронные книги революционны в неполитическом смысле, и тем не менее они могут изменить мир так же, как и пресс Гутенберга. Проект Гутенберг направлен на образование и грамотность, а также на обеспечение равных условий игры. Я не пытаюсь изменить людей; я пытаюсь изменить информационную инфраструктуру. Мы делаем электронные книги и раздаём их. Что люди с ними делают, полностью зависит от них».

Вот ещё несколько высказываний Харта:

– Я, конечно, неизлечимый романтик, и я верю, что технология может принести прогресс.

– Я делаю Проект Гутенберг по самым эгоистичным причинам – потому что хочу, чтобы в мире был Проект Гутенберг.

– Гутенберг не изобрёл печатный станок; он просто добавил один элемент: подвижный шрифт. Я не изобретал интернет, я просто добавил один элемент: книги, которые можно копировать со скоростью света. Как и Гутенберг, я сделал книги намного более доступными для масс и намного дешевле. Я был, безусловно, первым, кто сказал, что компьютеры и интернет будут использоваться не для вычислений, а для распространения общей информации. Но потребовалось много времени, чтобы идея прижилась, какое-то время никто особо не обращал внимания.

– Как бы Вы себя описали? – Идеалист, прагматик, трудоголик, новаторский, независимый, умный, предприимчивый, общительный. Короче говоря, я человек, претворяющий в жизнь свои мысли и идеалы; провидец, если хотите.

Ему удалось увидеть дело своей жизни реализованным. 50 лет назад он мечтал о 10 тысячах оцифрованных книг и первые годы работал в одиночку. К 1987 году ему удалось напечатать 313 книг, и этот период он называл борьбой с ветряными мельницами. Дело пошло быстрее, когда к проекту присоединились добровольные помощники из университетской группы пользователей компьютеров (это был один из первых примеров краудсорсинга). Очень своевременно появились концепция Всемирной паутины (Тим Бернерс-Ли, 1989) и браузер Mosaic (Марк Эндриссен, 1993). Благодаря растущей популярности интернета число волонтеров в 1990-е годы заметно увеличилось. Рубеж в 10000 книг был достигнут в октябре 2003 года. Юбилейный номер получила знаменитая Magna Carta (1215) – «Великая хартия вольностей», первый английский конституционный текст¹⁴.

Отметив 40-летие своего проекта, 64-летний Харт писал друзьям 16 июля 2011 г.: «Я достиг всех целей, которые ставил перед собой на протяжении всей жизни, и могу сказать, не опасаясь излишних резонансных откликов, что карьера, которую я выбрал в области электронных книг, оказалась успешной с точки зрения того, чего я пытался достичь за последние четыре десятилетия»¹⁵. И в том же письме, ясно осознавая ограниченность своих ресурсов и оставшегося времени,

¹⁴ <https://www.gutenberg.org/files/10000/10000.txt>

¹⁵ <http://brewster.kahle.org/2011/09/07/michael-hart-of-project-gutenberg-passes>

констатируя снижение уровня энергии, он ставит перед собой и перед проектом новые задачи. «Я хотел бы напомнить вам о своих последних целях. В свободном доступе около 25 миллионов книг. Если мы сделаем 40% из них, это будет 10 миллионов электронных книг. Существует около 250 языков, на которых говорят более миллиона человек. Если мы сделаем 40% из них, это будет 100 языков. Перевод 10 миллионов электронных книг на 100 языков даёт ОДИН МИЛЛИАРД ЭЛЕКТРОННЫХ КНИГ». Эта цель и сейчас выглядит фантастичной. До оцифровки 10 миллионов книг очень далеко; столь же утопичны перспективы автоматического перевода приемлемого качества. Но полвека назад и 10000 электронных книг казались чем-то нереальным. Харт любил повторять слова Конфуция: «путь длиной в тысячу ли начинается с первого шага».

Оставив работу по непосредственной оцифровке текстов, он «хотел бы как можно больше просто работать за кулисами. Я понимаю, что моё имя может чего-то стоить в связях с общественностью, я всё ещё способен писать много информационных бюллетеней, хотя мне нужно сделать гораздо больше в других аспектах моей жизни. Возможно, я смогу покинуть эту сцену без особых проблем, и, возможно, я смогу даже управлять ею заочно».

Письмо Харта воспринимается как его завещание. Жить ему оставалось меньше двух месяцев. Он умер 6 сентября 2011 г. от сердечного приступа в своём доме в Урбане. Проект к тому времени насчитывал более 36 тысяч книг.

В последние годы жизни Майкл Харт сосредоточился на организационных аспектах и популяризации проекта. С этой целью он в начале 2004 г. посетил Европу и выступил в ЮНЕСКО, на заседании Национальной ассамблеи Франции и на сессии Европейского парламента в Брюсселе. Однако ожидаемой реакции он не получил. «Мало надежды найти хоть какую-то поддержку в реальном мире. Я не думаю, что мир в целом действительно искренне хочет обеспечить грамотность и образование в странах третьего мира, несмотря на все разговоры об обратном. Поэтому существует вероятность того, что этот проект не будет поддерживаться из внешних источников».

Харт также был вынужден заниматься судебной защитой проекта. В интервью 2002 года он говорит: «Я только что закончил многолетнюю работу над отчётом для Верховного суда в надежде отменить последние расширения авторских прав». По «закону Сонни Боно» миллион потенциальных электронных изданий

был выведен из общественного достояния. Харт резонно начал опасаться за судьбу дела его жизни. Когда этот закон решили оспорить в Верховном суде, он согласился стать основным истцом, считая, что ему был причинён ущерб от продления срока действия авторских прав и нарушения конституции. Однако коллеги отказались приложить к судебным документам его эмоциональное заявление о том, что «закон аннулировал общественный договор, а общественное достояние насилуют и грабят», сочтя его чересчур популистским. В роли истца Харта заменил экс-программист Эрик Элдред, создатель электронной библиотеки Eldritch Press. Его поддержали Фонд свободного программного обеспечения, Национальный союз писателей, архивы, библиотеки, экономисты, в том числе пять нобелевских лауреатов. На стороне ответчика – правительства – выступило мощное лобби: крупные производители интеллектуального продукта, владельцы авторских прав, конгрессмены. Иск удовлетворён не был. Но, несмотря на проигрыш, инициаторам процесса удалось оказать влияние на общественное мнение, что в итоге привело к появлению лицензий Creative Commons¹⁶.

Вернёмся к интервью 2006 года.

«– Какие проблемы возникают с авторским правом в отношении работы Проекта Гутенберг?

– Самая большая проблема – это время, необходимое для проведения необходимого исследования авторских прав перед созданием электронного текста. Следующая проблема связана с постоянной угрозой судебных исков, которые возникают каждый год. Но самая большая проблема авторского права, с которой мы когда-либо сталкивались, возникла после того, как мы потратили почти все 1980-е, работая над изданием Шекспира. Ожидалось, что оно станет общественным достоянием, но законы об авторском праве были изменены.

– Как могло случиться, что Вам не удалось опубликовать книгу пьес автора, умершего 390 лет назад?

– Потому что издатель может получить авторские права на новое издание Шекспира, по-другому скомпоновав отрывки из предыдущих изданий».

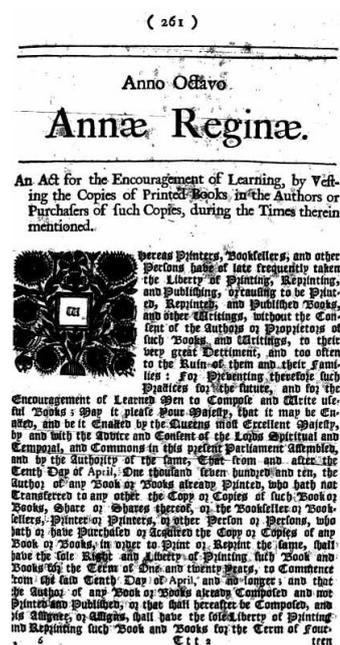
Главная проблема, обозначенная Майклом Хартом, требует отдельного рассмотрения.

¹⁶ <http://www.artpragmatica.ru/lessig/lessig.pdf>

ОБ АВТОРСКИХ ПРАВАХ

Законодательство США об авторском праве следует британскому Статуту королевы Анны¹⁷ (1709), который вступил в силу 10 апреля 1710 года и считается первым полноценным законом об авторских правах. Его полное название «*An Act for the Encouragement of Learning, by vesting the Copies of Printed Books in the Authors or purchasers of such Copies, during the Times therein mentioned*» (Акт о поощрении учёности путём наделения авторов и покупателей правами на копирование печатных книг на указанный период времени). Статут королевы Анны устанавливал 14-летний срок действия копирайта для всех публикаций. Его можно было возобновить один раз при жизни автора, после чего произведение становилось общедоступным.

Статут королевы Анны был не первым шагом в развитии копирайта. Иоганн Гутенберг со своей Библией создал большую проблему для церкви, лишившейся значительных доходов, но сохранившей своё влияние и возможности духовной цензуры. Свои экономические интересы защищали и светские власти, контролируя выдачу разрешений на книгопечатание. Сочетание этих факторов привело к появлению жёсткого законодательства. Так, 13 января 1535 года во Франции по требованию католической церкви был принят закон, полностью запрещающий книгопечатание и использование печатных прессов под угрозой повешения. Результатом стали вывод типографий за границы Франции и рост контрабанды книг. В Англии 4 мая 1557 года, за полтора века до Статута королевы Анны, появился Статут королевы Марии, вводивший монопольное право Лондонской гильдии печатников на тиражирование ранее выпущенных произведений, прошедших цензуру. Новая издательская монополия получила название «копирайт». Статут королевы Анны возник в результате борьбы авторов за свои интересы.



¹⁷ <http://www.copyrighthistory.com/anne.html>

Важное для демократии и современной цивилизации историческое событие произошло в Лондоне 22 февраля 1774 года. Палата лордов 22 голосами против 11 поддержала шотландского книготорговца Александра Дональдсона в его многолетней борьбе с книготорговой корпорацией. Ему удалось доказать, что сохранение бессрочной монополии противоречит британскому праву, а затруднение человечеству доступа к творениям гениев – здравому смыслу. Лорды с ним согласились, создав тем самым понятие общественного достояния. По новым законам по истечении определенного срока любое произведение становилось доступным для желающих его тиражировать¹⁸.

Первый федеральный закон США, регулирующий авторское право – Закон об авторском праве от 31 мая 1790 года, подписанный Джорджем Вашингтоном (*Copyright Act of 1790*)¹⁹, – унаследовал британские нормы. Как и Статут королевы Анны, он обеспечивает автору «исключительное право на свободу печати, переиздания и торговли», защищает оригинальные авторские произведения, зафиксированные на материальном носителе.

Интересно, что одним из влиятельных противников копирайта был Томас Джефферсон, третий президент США и один из авторов той самой Декларации независимости, с которой начался Проект Гутенберг. Он писал: «Если природа создала что-то менее пригодное для частной собственности, чем всё остальное, так это акт мыслительной силы, под названием идея, которой человек может обладать исключительно лишь до тех пор, пока он приберегает её для себя; но в тот самый момент, когда она оглашена, она вторгается в обладание каждого, и получивший её не может отказаться от обладания ею ... идеи должны свободно распространяться от одного к другому по всему миру во имя морального и взаимного наставления человека и улучшения его благосостояния»²⁰.

В течение XX века в США было три продления авторских прав. Как указал М. Харт в интервью 2006 года, в 1909 году это было сделано для противодействия

¹⁸ <https://inosmi.ru/europe/20120214/185848612.html>

¹⁹ <https://www.copyright.gov/history/1790act.pdf>, <https://www.copyright.gov/about/1790-copyright-act.html>

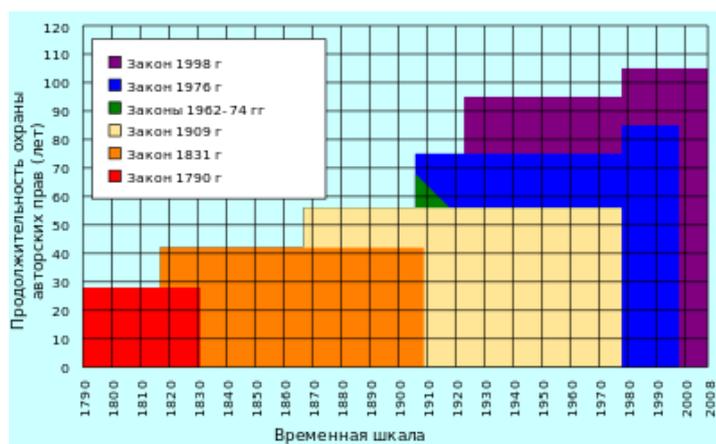
²⁰ <https://nauchkor.ru/uploads/documents/583e06b75f1be77312a3fd66.pdf>

паровым и электрическим прессам; в 1976 году – для противодействия копирувальным аппаратам Xerox; а в 1998 году – для противодействия интернету²¹.

Согласно закону 1976 года, авторское право сохранялось в течение жизни автора и в течение 50 лет после его смерти (для работ, сделанных по найму, – 75 лет). Закон 1998 года увеличил эти сроки до действия в течение жизни автора и 70 лет после его смерти, а для работ по найму – до 120 лет после создания или 95 лет после публикации в зависимости от того, какая дата раньше. Этот закон, подписанный президентом Б. Клинтоном и вступивший в силу 27 октября 1998 г., имеет номер Pub.L. 105–298, официальное название «Закон о продлении срока охраны авторских прав имени Сонни Боно» (*Copyright Term Extension Act, CTEA*), а в народе известен как «Закон о защите Микки Мауса»²². Расширив авторское право в США ещё на 20 лет, он тем самым вывел из общественного достояния миллион потенциальных электронных изданий, подпадающих под действие старых правил авторского права. В частности, Микки Маус, днём рождения которого считается 18 ноября 1928 года, станет общественным достоянием не ранее 2024 года. График отображает этапы увеличения срока авторского права в США²³.

Закон усложнил правила длительности авторских прав для произведений, созданных до 1978 года. При этом произведения, созданные

до 1923 года, продолжают оставаться в общественном достоянии. Помимо истечения срока защиты авторского права, произведение может оказаться общедоступным, если владелец прав сам передал его в общественное достояние. В общественном достоянии также находятся служебные произведения, созданные правительственными служащими США в рамках выполнения своих официальных обязанностей.



²¹ https://ia802307.us.archive.org/23/items/The_Basement_Interviews/Michael_Hart_Interview.pdf

²² https://ru.wikipedia.org/wiki/Copyright_Term_Extension_Act

²³ https://ru.wikipedia.org/wiki/История_авторского_права_в_США

Следует также упомянуть «закон о защите авторских прав в цифровую эпоху» (*Digital Millennium Copyright Act, DMCA*) номер Pub.L. 105-304²⁴, подписанный Б. Клинтонем и вступивший в силу 28 октября 1998 года. Его цель – упорядочение правового регулирования использования интернета. DMCA внёс поправки в раздел 17 Кодекса Соединённых Штатов, чтобы расширить сферу действия авторского права, ограничив при этом ответственность поставщиков онлайн-услуг за нарушение авторских прав их пользователями.

Майкл Харт прокомментировал препятствия, возникшие из-за новых ограничительных законов: «Мы сталкиваемся с ситуацией, когда рост доли информации, защищенной авторским правом, усугубляется постоянным продлением срока действия авторского права. Более 90% всех когда-либо предоставленных авторских прав всё ещё в силе. Другими словами, общественное достояние, которое составляло около 50% всего, что было написано столетие назад, упадёт примерно до 0.00001% или меньше на столетие вперед ... Средняя продолжительность авторского права сейчас составляет 95.5 лет, что резко противоречит первоначальной идее 14-летнего авторского права, созданной Статутом Анны»²⁵.

Большинство книг из коллекции проекта Гутенберга распространяется как общественное достояние в соответствии с законодательством США. Есть также несколько защищённых авторским правом текстов, которые проект распространяет с разрешения авторов, например, тексты фантаста Кори Доктороу. На них распространяются дополнительные ограничения, установленные правообладателем, хотя обычно они лицензируются в соответствии с Creative Commons. Американская некоммерческая организация Creative Commons, созданная в декабре 2001 г., опубликовала для бесплатного использования несколько лицензий, с помощью которых авторы могут гибко управлять своими правами. Эти лицензии Creative Commons позволяют авторам-создателям сообщить, от каких прав они хотели бы отказаться, – например, автор не хочет запрещать пользователям распространять своё произведение (по закону такой запрет включается автоматически при создании произведения), – а какие права они оставляют за собой. Этот

²⁴ <https://www.copyright.gov/legislation/dmca.pdf>

²⁵ https://ia802307.us.archive.org/23/items/The_Basement_Interviews/Michael_Hart_Interview.pdf

подход назван «some rights reserved» («некоторые права защищены»), в противовес стандартному «all rights reserved, все права защищены». Лицензии Creative Commons юридически не противоречат и не заменяют авторское право, наоборот, основаны на нём и используют его законы, но из-за более либерального подхода создают конкуренцию стандартному ограничительному авторскому праву и воспринимаются как оппозиция копирайту²⁶.

В нашей стране произведение переходит в общественное достояние, как правило, через 70 лет после смерти автора. Федеральным законом от 12 марта 2014 года № 35-ФЗ «О внесении изменений в часть первую, вторую и четвёртую Гражданского кодекса Российской Федерации и отдельные законодательные акты Российской Федерации» были приняты поправки, которые коснулись защиты интеллектуальных прав²⁷. В частности, введена новая статья 1286-1 об открытой лицензии на использование произведения науки, литературы или искусства. Предусмотрено, что лицензионный договор, по которому автором (иным правообладателем) предоставляется простая (неисключительная) лицензия на использование произведения, может быть заключён в упрощённом порядке (открытая лицензия).

ПРОЕКТ ГУТЕНБЕРГ. ФИЛОСОФИЯ, ИСТОРИЯ, РАЗВИТИЕ, СТАТИСТИКА

Девиз Харта и его проекта – «Сломать барьеры невежества и безграмотности!». Харт стремился предоставить «как можно больше электронных книг как можно большему количеству людей» и сформулировал миссию проекта: «*Поощрять создание и распространение электронных книг*»²⁸.



Сейчас электронные книги стали глобальным явлением; люди читают книги на своих телефонах, гаджетах и других устройствах. А полвека назад это было редкой экзотикой, и до середины 1990-х критики проекта считали его реализацию в заявленных масштабах невозможной.

²⁶ <https://creativecommons.org/licenses>

²⁷ <http://www.kremlin.ru/acts/bank/38195>

²⁸ <https://www.gutenberg.org/help/faq.html>

В документе «История и философия Проекта Гутенберг» [8], датированном августом 1992 г., Майкл Харт изложил основные принципы, которые он положил в основу проекта и которым неуклонно следовал. Предпосылка, на которой строился проект, заключалась в следующем: всё, что можно ввести в компьютер, можно воспроизводить бесконечно. Это Майкл назвал «технологией репликатора» (Replicator Technology). Концепция этой технологии проста; как только книга или любой другой объект (включая картинки, звуки и даже трёхмерные изображения) могут быть сохранены в компьютере, любое количество копий может и будет доступно, и каждый сможет иметь копию книги, которая была введена в компьютер. Там же говорится: «Библиотека Проекта Гутенберг состоит из трёх разделов:

- лёгкая литература – такая как «Алиса в стране чудес», «Алиса в Зазеркалье», «Питер Пен», басни Эзопа и т. д.;
- серьёзная литература – Библия и другие религиозные труды, Шекспир, «Моби Дик», «Потерянный рай» и т. д.;
- справочники и рекомендации – тезаурусы, альманахи, энциклопедии, словари и т. д.»

Философия Проекта Гутенберг заключается в том, чтобы сделать информацию, книги и другие материалы доступными для широкой публики в формах, которые подавляющее большинство компьютеров, программ и людей может легко читать, использовать, цитировать и искать. Как следствие, электронные тексты должны стоить настолько дёшево, чтобы никого не волновало, сколько они стоят. Их размеры должны подходить для стандартных носителей своего времени.

К этому следует относиться серьёзно. Харт настаивал на преимущественном использовании ASCII (American Standard Code for Information Interchange), нижнего уровня американского стандарта кодирования для обмена информацией, распознаваемого в 99% программного обеспечения. Это требование может показаться мелочью, даже старомодностью, но на самом деле оно очень важно. С 1980-х годов появлялось множество соблазнов «обновиться»: такие форматы файлов как WordStar, TeX или LaTeX в 1980-х, или XyWrite, MS Word или Adobe Acrobat в 1990-е и 2000-е годы предоставляли более широкие возможности форматирования, чем ASCII: курсив, полужирный шрифт, позиции табуляции, выбор

шрифтов, сноски, формулы, представление страниц и т. д. Но опытный программист Харт понимал, что «оптимальные форматы» часто меняются, а чтение электронного текста в формате ASCII через ftp или через веб-браузер не требует изменения программного обеспечения, тем самым обеспечивая максимально широкий круг читателей. Только предельно лёгкие и дешёвые в доступе электронные книги проекта могли быть распространены на большую аудиторию²⁹.

Развитие технической базы позволило смягчить эту политику. Книги стали сопровождаться иллюстрациями. В дополнение к ASCII-файлам, хранящимся на перфолентах или магнитных лентах, постепенно появились файлы в форматах html и pdf, а позднее – в популярных форматах электронных книг, в частности, ePUB. Помимо книг, проект получил такие дополнительные типы контента, как музыкальные партитуры, фильмы и звуковые файлы. В сентябре 2003 г. были выпущены электронные книги Project Gutenberg Audio. Аудиокниги, читаемые людьми, полезны для слабовидящих читателей. В том же году появились копии книг на CD и DVD. Компакт-диск «Best of Gutenberg» от августа 2003 года содержал более 600 книг, а в декабре вышел первый DVD «Project Gutenberg» с 9400 наименований. Однако основой проекта остаются традиционные книги.

Харт предпочитал тексты, которые представляли собой исторический интерес и были достаточно небольшими, чтобы поместиться на одной дискете. Во-первых, их было легко печатать; во-вторых, в те дни для многих компьютерных пользователей единственным способом передачи файлов была отправка дискет по почте. Когда проект только начинался, стандартная книга на 300 страниц занимала один мегабайт, чего в 1971 году, как правило, не было ни у одного пользователя. Носителями информации были 8-дюймовые дискеты ёмкостью 80 Кбайт, поэтому 5-килобайтная Декларация независимости была хорошим началом. Затем появилась возможность оцифровывать отдельные главы Библии и Шекспира (пьеса за пьесой). В начале 1980-х годов, когда Проект Гутенберг уже был широко известен, стандартом стали 5-дюймовые дискеты по 360 Кбайт, и книги вроде «Алиса в стране чудес» и «Питер Пэн» стали помещаться на одной дискете. В середине 1980-х годов размер дискет уменьшился до 3.5 дюймов, что при объёме

²⁹ <https://www.chronicle.com/article/michael-hart-1947-2011-defined-the-landscape-of-digital-publishing>

1.44 Мбайт и стандартном zip-сжатии позволило довести размер файла практически до 3 миллионов знаков, чего достаточно для большинства книг³⁰.

Как мы видели, в течение 17 лет Харт вёл свой проект в одиночку. Фактически за 20 лет до широкого распространения интернета он стал первым информационным провайдером, когда в сети было едва ли 100 человек. Однако мало кто проявлял интерес к тому, что он делал; его называли «чокнутым парнем, который хочет ввести всего Шекспира в компьютер». Много изменилось, по словам Харта, примерно в 1988–1989 годах, когда он попал на собрание университетской группы пользователей ПК, после чего возник кружок помощников-энтузиастов. Там же присутствовал Марк Зинцов (Mark Zinzow) – старший программист-исследователь, который оказал огромную помощь в организации работы зеркальных сайтов Проекта Гутенберг, списков рассылки, почтовых серверов. Благодаря этому вскоре начали появляться волонтёры. Уникальная для своего времени особенность проекта состоит в том, что он стал одним из первых примеров использования сети для совместной работы группы географически рассредоточенных добровольцев над общей задачей (такую же модель впоследствии использовало движение за открытый исходный код Open Source movement). К тому времени интернет насчитывал четверть миллиона пользователей, а добровольцев было, по оценке Харта, порядка 10 тысяч человек.

Один из важнейших вопросов для электронных библиотек – адекватность представления текста. Вспоминается, как лет 20 назад в МГУ в ходе практикума по поиску в интернете мои студенты обнаружили три различных варианта последней фразы рассказа «Каштанка». В сети случалось видеть «Войну и мир» без французских фрагментов, урезанные описания природы. Сам Харт писал по этому поводу: «Проект является чисто волонтёрским и не стремится к перфекционизму. Мы не пишем для читателя, которого волнует, стоит ли у Шекспира двоеточие или точка с запятой между словами. Мы нацелены на выпуск текстов с точностью 99.9% в глазах обычного читателя. Учитывая предпочтения корректоров и общее отсутствие у публики навыков чтения, мы, вероятно, значительно превышаем эти требования. Однако тому, кто хочет получить «авторитетное издание», придётся

³⁰ <https://www.elibrary.ru/item.asp?id=26739120>

подождать некоторое время». Как бы то ни было, тщательная вычитка текста – важная часть процесса.

В 1990-е годы началось сканирование текстов с использованием технологии оптического распознавания символов (optical character recognition, OCR). Появился веб-сайт проекта. Качественный прорыв наступил после подключения к нему движения «Распределённые корректоры» (Distributed Proofreaders, DP). Как сказано на сайте движения³¹, задача этой общественной инициативы – «сохранение текстов, которые находятся в общественном достоянии, путём преобразования их в высококачественные, бесплатные цифровые транскрипции благодаря усилиям поддерживающего сообщества добровольцев». DP были основаны в конце 2000 года Чарльзом Фрэнксом (Charles Franks) и вначале задумывались для оказания помощи Проекту Гутенберг, но вскоре стали основным источником электронных книг проекта. В 2006 г. Distributed Proofreaders стал отдельным юридическим лицом, продолжая поддерживать тесные отношения с Проектом Гутенберг.

Процесс происходит так. Менеджеры проектов сканируют книги, а полученные изображения обрабатывают с помощью OCR. Но при распознавании текста часто появляются ошибки. Для их исправления отдельные страницы через веб-интерфейс получают волонтеры-корректоры, при этом они видят одновременно и изображение исходной страницы, и распознанный текст. Добровольцы вносят изменения в текст, сохраняют его и возвращают менеджеру, который устраняет проблемы, отмеченные корректорами, а затем отправляет электронный текст в Проект Гутенберг. Каждое издание вычитывают два разных человека. Distributed Proofreaders позволяет нескольким корректорам работать над одной книгой одновременно. Считается, что после OCR текст в лучшем случае надёжен на 99%. Ручная корректура доводит это значение до 99.95%, аналогичный стандарт принят в Библиотеке Конгресса США.

Трудоёмкий процесс исправления ошибок эффективно распределяется аналогично распределённым вычислениям. Ежемесячно электронная библиотека пополняется на 150–200 книг, причём ежедневно вычитку проходят более 2 тысяч страниц.

³¹ <http://www.pgdp.net/c/default.php>

В 1994 году Проект Гутенберг впервые стал самофинансирующимся, собрав достаточно пожертвований, чтобы покрыть все расходы. К тому времени Харт потратил на проект около \$100 тысяч из личных средств.

Помимо «Распределённых корректоров», у проекта есть другие партнёрские организации. Харт, как мы видели выше, занимался организационной деятельностью. В 2000 году была образована общественная благотворительная организация Фонд литературного архива Проекта Гутенберг (The Project Gutenberg Literary Archive Foundation, PGLAF). Фонд, как сказано на его сайте, управляет проектом³². 19 марта 2002 года Майклом Хартом зарегистрирована торговая марка «Project Gutenberg». Университет Карнеги-Меллона по соглашению с Хартом взял на себя управление финансами проекта. С 2003 года с Проектом Гутенберг связаны публичная библиотека iBiblio и Консорциумный центр (PGCC). На сайте проекта приведён список некоторых партнёров³³. На ряде других ресурсов популяризуется бесплатная или недорогая литература и распространяются книги Проекта Гутенберг.

В 2000-х «дочерние» проекты вышли за пределы США. Начало положил Project Gutenberg-DE³⁴, сосредоточившийся на немецкой литературе. Он предоставляет бесплатные электронные книги на веб-сайте, CD и DVD. На Project Gutenberg of Australia³⁵ организован свободный доступ к книгам австралийского филиала, а также текстам оригинального проекта, которые написали австралийцы и исследователи континента. Важно, что в Австралии закон о защите авторских прав более либерален, чем в США; это расширяет круг доступных произведений. Project Runeberg³⁶ занимается литературой скандинавских стран. Специализация Project Gutenberg of Canada³⁷ – канадская литература на английском и французском языках. Аналогичные проекты действуют в таких странах, как Финляндия, Люксембург, Филиппины, а также Тайвань. Всё это совершенно разные организа-

³² <http://cand.pglaf.org>

³³ https://www.gutenberg.org/about/partners_affiliates.html

³⁴ <https://www.projekt-gutenberg.org>

³⁵ <http://gutenberg.net.au>

³⁶ <http://runeberg.org>

³⁷ <http://gutenberg.ca>

ции. Как правило, Харт давал им разрешение на использование зарегистрированной торговой марки «Project Gutenberg». Эти сайты действуют в рамках норм авторского права своих стран и направлены на определённые типы национальных произведений и языков³⁸. Проект Гутенберг также сотрудничает с негосударственным культурно-просветительским проектом Растко³⁹, основанным в 1997 году. Он назван в честь сербского просветителя Растко Неманича (1174–1256) и как часть балканской культурной сети опирается на сеть из двух десятков региональных центров в восточной Европе и на Балканах; оказывает поддержку культуре этнических меньшинств. На Растко поддерживается европейская версия DP, снабжающая Проект Гутенберг текстами общественного достояния на европейских языках.

Не имеет смысла оценивать размер Проекта Гутенберг в мега- и гигабайтах, ведь в зависимости от формата «вес» электронной книги изменяется. Аудиофайл минутной длительности занимает столько же места, как десятки книг в формате ASCII. Тем не менее, существуют оценки, показывающие, что по количеству слов Проект Гутенберг превосходит Александрийскую библиотеку⁴⁰. Сейчас размер проекта ограничен не столько работоспособностью добровольного сообщества, сколько доступностью материала из общественного достояния, не подпадающего под ограничения авторских прав.

Некорректно сравнивать Проект Гутенберг с Библиотекой Конгресса США: очевидно, что она на несколько порядков крупнее. Но, во-первых, Библиотека Конгресса содержит экземпляры всех книг, зарегистрированных с авторским правом в США, а во-вторых, её коллекция включает множество переизданий. Столь же неуместно сравнение с проектом Google Book Search, книжным сервисом IT-гиганта Google с его многомиллионными объёмами отсканированной печатной продукции, живущим во многом за счёт рекламы и платных подписок. Проект Гутенберг создаётся обществом в интересах общества. По словам Т.Е. Савицкой (РГБ), «значимость Проекта Гутенберг как первой краудфандинговой модели в сфере электронных библиотек заключается отнюдь не в объёме оцифрованной

³⁸ <https://www.gutenberg.org/help/faq.html>

³⁹ <http://www.rastko.rs>

⁴⁰ http://freesoftwaremagazine.com/articles/impossible_thing_2_comprehensive_free_knowledge_repositories_wikipedia_and_project_gutenberg

литературы, многократно превзойдённом как Google Book Search – частным проектом глобальной корпорации, так и, например, электронными библиотеками Internet Archive и Digital Public Library of America, имеющими поддержку благотворительных фондов. Уникальность Проекта Гутенберг в структурном воспроизводстве логики формирования цифрового универсума: предельной децентрализации как организационном принципе, спонтанном демократизме, опоре на сетевое сообщество, которые превратили его, фактически, в кластер родственных многоязычных проектов. Опыт такого рода строительства электронных библиотек, опирающихся на интерактивный потенциал глобального общества Web 2.0, чрезвычайно важен и должен пристально изучаться»⁴¹.

Приведём данные о росте книжной коллекции проекта (они не включают сведения о сайтах Австралии, Канады и др.). Таблицы составлены по материалам сайта Проекта Гутенберг⁴².

Таблица 1. Динамика роста Проекта Гутенберг

Год	Число книг
1997, август	1000
1999, май	2000
2000, декабрь	3000
2001, октябрь	4000
2002, апрель	5000
2003, октябрь	10000
2006, декабрь	20000
2009, сентябрь	30000
2012, июнь	40000
2015, сентябрь	50000
2019, июль	60000

«Юбилейные» номера обычно присваивались наиболее известным и значимым текстам:

⁴¹ <https://bibliotekovedenie.rsl.ru/jour/article/viewFile/534/480>

⁴² <https://www.gutenberg.org>

Таблица 2. Номера некоторых публикаций

10	Библия короля Якова (1769)
100	Полное собрание сочинений Шекспира (1590-1613)
1000	Божественная комедия Данте (1321, на итальянском яз.)
2000	Дон Кихот Сервантеса (1605, на испанском яз.)
3000	Под сенью девушек в цвету Пруста (1919, на французском яз.)
5000	Записки Леонардо да Винчи
10000	Великая хартия вольностей
20000	Двадцать тысяч лье под водой Жюль Верна
50000	Биография Гутенберга

Отметим, что номер 1984 зарезервирован для одноимённого романа Дж. Оруэлла (текст готов к размещению, но подпадает под действие авторского права и ждёт публикации до 2045 г.).

К осени 2021 г. Проект Гутенберг насчитывал более 66 тысяч книг. Из них 50 тысяч на английском языке; заметно отстают французский (3000), финский (2100), немецкий (1800), испанский (650), португальский (550). Русский язык представлен десятком названий; из них наибольший интерес представляют, пожалуй, две повести Достоевского и «Детство» Толстого. Есть ещё три сотни переводов русских авторов – в основном на английский и почему-то финский; значительно меньше – на немецкий, французский и эсперанто⁴³. Самое популярное произведение – «Франкенштейн» Мери Шелли (104 тысячи загрузок).

О ХАРТЕ. ПОСЛЕ ХАРТА

Знакомые отмечали особенности характера Харта; некоторые считали его эксцентричным чудаком. По этому поводу он высказался в интервью 2006 года. «Я специально избегаю термина «эксцентричный». Тем не менее, я признаю, что не похож на кого-либо, кого я когда-либо встречал. Я думаю не так, как другие, иначе вы бы со мной не разговаривали. С другой стороны, если бы вы встретили меня где-нибудь, я бы не выделялся, разве что я готов разговаривать практически

⁴³http://www.gutenberg.org/ebooks/search/?query=Russian+&submit_search=Go%21&start_index=301

со всеми. Не думаю, что я выделялся среди своих сверстников в каком-либо отношении, и дело не в том, как я одевался или вёл себя. Просто мои интересы отличались от их, но большинство никогда этого не осознавало». Его необычный взгляд на мир отмечал заместитель декана Иллинойского университета: «Существует опасность, что Харта, как и многих других блестящих молодых людей, ошибочно сочтут просто эксцентричным. Поэтому ему должна быть предоставлена возможность придерживаться собственного курса действий, который на первый взгляд может показаться несколько странным». Вероятно, родители Харта, профессора Иллинойского университета, воспитали в нём стремление искать истину и подвергать сомнению авторитеты – черты настоящего интеллектуала.

Лучшим свидетельством решимости Харта – человека, которого нелегко отклонить от своей цели, является прежде всего Проект Гутенберг. Как выразилась его знакомая Сью ДеВрис, «Майкл очень храбрый в полном смысле слова, что позволяет ему придерживаться избранного пути, когда другие могут счесть задачу слишком сложной. Он не меняет своего мнения под влиянием общественного мнения или внешнего давления. К тому же он очень забавный, наслаждается многими вещами и здорово смеётся».

В глазах многих Харт был героем благодаря его неуклонной приверженности открытым стандартам и свободному доступу к литературе; отстаиванию философии, согласно которой историческое наследие человечества принадлежит самому человечеству. Не будем забывать, что, когда Харт формулировал основополагающие принципы Проекта Гутенберг, принцип открытости был крайне радикальным. Вплоть до конца 1990-х парадигма издательского дела строилась на презумпции права собственности. Даже если в Декларации независимости говорилось, что мы обладаем «неотъемлемыми правами», право читать эту Декларацию в значительной степени ограничивалось теми, у кого были читательские билеты библиотек или кто мог купить у издателя её копию.

Трудности общения с Хартом отмечал и интервьюер Р. Пойндер. «Как я обнаружил, Харт – сложный человек для интервью. Разговор с ним – это, как он сам выражается, «пасти кошек». <Эта идиома (herd cats)⁴⁴, обозначающая бессмысленное занятие, стала популярной после выхода в январе 2000 года рекламного

⁴⁴ https://en.wiktionary.org/wiki/herd_cats#Etymology

ролика компании Electronic Data Systems; впрочем, ещё в 1979 г. она встречалась в комедии группы Монти Пайтон «Житие Брайана» (Monty Python's The Life of Brian)⁴⁵ – ЮП». Почти каждый задаваемый ему вопрос разбивался на серию связанных и более сложных вопросов, а затем возвращался ко мне. Поскольку предпочитаемый способ связи Харта – электронная почта, мой почтовый ящик вскоре заполнился сотнями сообщений; и когда он, наконец, согласился на телефонный разговор, он настоял, чтобы я выключил магнитофон ... Он также кажется человеком, внимание которого может быстро отвлекаться. Мне становилось все труднее заставить Харта ответить на мои просьбы прояснить какие-то моменты: его постоянно активный ум, казалось, переключался на новые схемы и идеи»⁴⁶.

4 июля 2011 Мари Лебер, автор книги [9], подготовила иллюстрированный альбом «The Project Gutenberg EBook of Project Gutenberg 4 July 1971 – 4 July 2011»⁴⁷, где перечислила основные вехи истории проекта за 40 лет его существования. Альбом она посвятила всем волонтерам Проекта Гутенберг и Distributed Proofreaders по всему миру.

На сайте Иллинойского университета помещён обширный список ссылок⁴⁸ на литературу о проекте (естественно, англоязычную).

Вероятно, первое упоминание о проекте на русском языке содержится в книге [5]. Там на с. 105 указано, что «пользователям доступны произведения Шекспира, Киплинга, Кэрролла, тексты Библии, Корана, Декларации независимости США, первый миллион цифр числа π и т. д.». Кстати, именно стихотворение Киплинга в оригинале, понадобившееся автору много лет назад для презентации, Проект Гутенберг быстро и чётко предоставил.

Харт не скрывал удовольствия, получая такие электронные письма: «Проект Гутенберг – это отличная штука. Я очень люблю читать. И только в «Гутенберге» я встретил пять книг, которых не мог найти ни в одном магазине или библиотеке»⁴⁹. Он говорил по этому поводу: «Люди просто читают, читают и читают, и они счастливы от этого». В 1999 г. газета Chicago Tribune напечатала его слова: «В мире есть

⁴⁵ http://montypython.50webs.com/scripts/Life_of_Brian/1.htm

⁴⁶ <https://poynder.blogspot.com/2006/03/interview-with-michael-hart.html>

⁴⁷ <https://www.gutenberg.org/files/36616/36616-pdf.pdf>

⁴⁸ <http://50years.ischool.illinois.edu/bibliography/gutenberg.html>

⁴⁹ <https://will.illinois.edu/news/story/e-book-developer-michael-hart-dies>

две абсолютно бесплатные и неиссякаемые вещи: это воздух, которым мы дышим, и электронные тексты из Проекта Гутенберг»⁵⁰. Одним из его любимых занятий было рассуждение о будущем, и Харт верил, что со временем любой желающий сможет иметь полную собственную копию библиотеки Проекта Гутенберг.

Харту нравилось высказывание Бернарда Шоу (George Bernard Shaw): «Разумные люди приспособляются к миру. Неразумные люди пытаются приспособить мир к себе. Поэтому весь прогресс зависит только от людей неразумных»⁵¹. Сам Харт, безусловно, относился к «неразумным», и мы во многом обязаны ему колоссальным прогрессом в создании и использовании электронных библиотек. Его «долг» в 100 миллионов долларов погашен с лихвой.

БИБЛИОТЕКА МОШКОВА



Очень много общего с проектом Харта у старейшей электронной библиотеки рунета, известной по имени своего создателя Максима Мошкова⁵². Её днём рождения считается 1 ноября 1994 года, когда ещё не было ни Яндекса, ни Рамблера. Мошков говорит: «Не помню точно, когда открылась библиотека, – это было в ноябре. Если народ считает днём начала работы сайта 1 ноября, что ж, я не против»⁵³. С «Гутенбергом»

её роднит то, что при очень скромном бюджете она держится в основном на энтузиазме основателя. Другие общие черты: подчёркнутая забота об удобстве пользователей, не изменяющийся годами (и даже десятилетиями) формат представления текстов, предельно лаконичный дизайн, а также неприязнь к законам об авторских правах. Библиотека Мошкова живёт согласно изречению из китайской Книги перемен: «Если менять не нужно, менять не надо», и до сих пор существует в первоначальном дизайне. Много лет назад её создатель сформулировал

⁵⁰ https://distributedmuseum.illinois.edu/exhibit/michael_hart/

⁵¹ Сравните: «Не стоит прогибаться под изменчивый мир, / Пусть лучше он прогнётся под нас». (Макаревич А.В., альбом «Отрываясь» (1997); https://ru.wikipedia.org/wiki/Однажды_мир_прогнётся_под_нас).

⁵² <http://lib.ru>

⁵³ <https://d-russia.ru/biblioteke-moshkova-ispolnyaetsya-22-goda.html>

«правила хорошего тона в WWW»⁵⁴, где, в частности, дал такой портрет: «посетитель сейчас, в 1996 году, живёт под Windows 3.11 на PC/P486 RAM 4M, и экран у него 14 дюймов, графическое разрешение 640×480 16 цветов. И модем у него 9600, шипящий по паршивой телефонной линии» (это выразительное описание полностью относится и к автору этого текста). Конечно, такую технику сейчас можно встретить разве что в музеях, но, как и в своё время Харт, Мошков стремится к тому, чтобы его библиотека была доступна пользователям даже с самыми скромными возможностями, и не изменяет то, что хорошо работает.

Максим Евгеньевич Мошков родился 13 октября 1966 года. Окончил механико-математический факультет МГУ. В 1991–2013 гг. работал в НИИ системных исследований РАН⁵⁵ (администратор институтской сети), много лет вёл ряд учебных курсов по Unix, HP Open View, HP DataProtector, VMware в образовательном центре HPE⁵⁶. Разработал программное обеспечение интернет-проектов Газета.ру, Лента.ру, Вести.ру, «Еженедельный журнал» (позднее «Ежедневный журнал») и др. Многократный лауреат различных интернет-премий. Автор книги «Введение в системное администрирование UNIX»⁵⁷. Указом №529 от 26.08.2020 награждён медалью ордена «За заслуги перед отечеством» II степени⁵⁸. Библиотека Мошкова в 2003 году стала лауреатом Национальной интернет-премии, в 2004 – Премии рунета. Более подробно о Мошкове можно прочесть на сайте библиотеки⁵⁹.

О зарождении и развитии библиотеки рассказывает её создатель. «Когда я кончил учиться и поступил на работу, то вдруг обнаружил, что есть компьютеры, в компьютерах есть файлы, и среди файлов встречаются файлы книжек. Их набивали, ими менялись. Я застал файлы, у которых была удивительная кодировка, путаница с большими и маленькими буквами. Я добывал файлы из интернета, из других мест, где водились электронные книги. Облазил все российские сервера, которые существовали на тот момент, их было, наверное, штук 15 в России и штук

⁵⁴ http://www.lib.ru/WEBMASTER/sowetywww.txt_with-big-pictures.html

⁵⁵ <https://www.niisi.ru>

⁵⁶ <https://education.hpe.com/ru/ru/training>

⁵⁷ https://intuit.ru/goods_store/ebooks/8382

⁵⁸ <http://publication.pravo.gov.ru/Document/View/0001202008290001?index=9>

⁵⁹ <http://lib.ru/~moshkov>

30 за рубежом. Всё тащил к себе, как белка в дупло, и укладывал. Потом туда потихонечку начал ходить народ»⁶⁰. «Вначале библиотека была просто средством отладки технологий интернета. Мне нужно было научиться делать сайты и наполнять их контентом. Где-то в 1990 г. я начал собирать электронные файлы. Если встречал интересную книгу, то копировал, оформлял и размещал на своём компьютере. Когда появился интернет, я эту коллекцию разместил в сети»⁶¹. «В 1994 году я работал в Отделении математики РАН, нас подключили к WWW, и в качестве тренировки каждый сделал по личной страничке. У меня там было про хобби — турпоходы и библиотеку, которую я собрал. Она потом и стала библиотекой Мошкова»⁶². «Когда пришёл работать в институт, обнаружил, что часть книг существует в цифровом формате ЭВМ. Понял, что это удобно, и собрал все файлы в один компьютер — так и началась электронная библиотека. В 1994 году она насчитывала всего двадцать томов»⁶³. «Делал для себя склад — хранилище текстов, заметок, документации. И закидывал в него всё, что мне было интересно. Оказалось, что эта коллекция интересна не только мне, но и другим. Так само собой и оказалось, что открывал персональную записную книжку, а открыл библиотеку»⁶⁴. Сначала туда попадало всё, что удавалось найти. Затем книги стали присылать читатели — кто-то сканировал, кто-то находил. А потом начали приходить авторы, которые сами предлагали книги в библиотеку. «Однако уже несколько лет я практически не пополняю библиотеку. Вместо этого я разработал технологию авторской самопубликации, на базе которой работает несколько сайтов. Самый крупный из них — система «Самиздат». Библиотека не заброшена, но живёт сама по себе — движок работает, книги доступны. Lib.ru — памятник эпохи, напоминает мне московский особнячок с собственным садиком и колодцем, уцелевший до наших дней»⁶⁵.

⁶⁰ <https://www.svoboda.org/a/24203692.html>

⁶¹ <http://www.unkniga.ru/innovation/tehnology/3379-maksim-moshkov-socialnye-seti-voruyut-avtorov-i-chitateley.html>

⁶² <https://www.colta.ru/articles/society/21190-maksim-moshkov-budet-belyy-internet-obsluzhivayuschiy-po-pasportam-i-chernyy-na-obochine>

⁶³ <https://yeltsin.ru/news/sozdatel-libru-maksim-moshkov-est-li-budushee-u-bibliotek>

⁶⁴ http://samlib.ru/r/romanow_w_e/pf3.shtml

⁶⁵ <https://d-russia.ru/biblioteke-moshkova-ispolnyaetsya-22-goda.html>

Этот запоминающийся адрес сейчас знаком миллионам пользователей. Но домен lib.ru появился только 22 декабря 1998 года⁶⁶, а до этого библиотека размещалась на сайтах <http://www.moshkow.orc.ru>, <http://kulichki.rambler.ru/moshkow>, <http://www.parkline.ru/Library>, <http://moshkov.relline.ru> и ещё на двузначном числе зеркал в России, Белоруссии, Эстонии, Нидерландах и на Украине [7].

Высококласный программист, Максим Мошков создал для собственной библиотеки десятки надёжных и работоспособных скриптов, обрабатывающих присылаемые текстовые файлы и поддерживающих файловую систему. Изначально он пользовался системой и текстовыми редакторами, разработанными в конце 1980-х годов. Соответственно, вся первая коллекция размещена в формате текстового редактора «РК». Тексты лежат в формате html, а не в Word (.doc): нажимая на заголовок той или иной работы, можно сразу же открыть её, никуда не скачивая. Довольно быстро библиотека превратилась из скромной и крайне хаотичной коллекции надёрганных «с миру по нитке» текстов в чётко структурированный ресурс, способный к самоподдержке и саморазвитию. При этом проявились визионерские качества Мошкова: задолго до появления Википедии он предвосхитил «вики-принцип»: пополнение библиотеки – задача в первую очередь посетителей. «Читатели стали присылать мне книги в виде файлов. Оказалось, чтобы создать библиотеку, не надо ни искать, ни выпрашивать. Она сама росла, по тому же принципу, по какому Том Сойер красил забор»⁶⁷.

Инструментом для такого пополнения служит уже упоминавшийся сервер современной литературы «Самиздат»⁶⁸, предназначенный для создания авторских литературных разделов. Как указано на сайте, его день рождения – 29 мая 2000 года. Говорит Мошков: «Я недаром заваял систему «Самиздат», чтобы в нашей русской Сети была куча художественной литературы заведомо бесплатной. Причём объёмы там такие, что уже сейчас хватит примерно на пять человеческих жизней. Хорошо читающий человек за свою жизнь прочитает 3 тыс. книг. И в «Самиздат» каждый день приходит штук пять новых толстых книг. Он пополняется

⁶⁶ <https://whois.ru/lib.ru>

⁶⁷ <https://bookmix.ru/news/index.phtml?id=12168>

⁶⁸ <http://samlib.ru>

быстрее, чем мы в состоянии прочитать. В «Самиздате» примерно полтора миллиона читателей в месяц. В русской классике – 800 тысяч. В библиотеке – ещё 800 тыс. читателей за месяц. Есть ещё мелкие дополнительные проекты – музыка и прочее. Если всё сложить, получится около 3 миллионов»⁶⁹. В 2021 году в журнале насчитывается свыше 105 тысяч авторов, приславших более полутора миллионов произведений.

Разумеется, авторы, присылающие свои материалы для публикации в интернете, не претендуют на защиту своих прав. Однако 12 марта 2014 г. появился федеральный закон №35-ФЗ «О внесении изменений в части первую, вторую и четвёртую Гражданского кодекса Российской Федерации и отдельные законодательные акты Российской Федерации»⁷⁰. По поводу 4-й части ГК президент РФ В.В. Фёдоров считает, что «у нас появился один из самых суровых из существующих в мире законов, охраняющий авторские права. С одной стороны, в нашей Конституции записано, что человеку даровано право на бесплатную информацию, а с другой стороны, что надо соблюдать права собственности, в том числе интеллектуальной. Этот баланс ищется. Сейчас чаша весов явно склонилась в пользу авторов»⁷¹. Позиция Максима Мошкова по этому поводу выглядит достаточно естественной. «Авторское право больше вредит культуре, чем помогает ей. Но помогает писателям и околоиздательским структурам. Кто из них более достоин — общечеловеческая культура или писатели? У меня есть мнение, но я не буду его озвучивать. Остап Бендер завещал нам чтить Уголовный кодекс, и деваться некуда»⁷². С Мошковым солидарен Илья Ларин, основатель электронной библиотеки «Либрусек»⁷³: «Авторское право в текущем варианте абсурдно. В идеале автор не должен иметь права запрещать использовать свои произведения, но должен получать некую разумную часть прибыли. Регистрация прав должна быть не автоматической, а по заявке автора, платной и на небольшой срок» [4].

⁶⁹http://www.chaskor.ru/article/maksim_moshkov_skolkovo_ya_vosprinimayu_kak_horoshij_znak_20738

⁷⁰ <http://www.kremlin.ru/acts/bank/38195>

⁷¹ <http://marc21.rsl.ru/index.php?doc=2911>

⁷² <https://godliteratury.ru/articles/2016/10/13/maksim-moshkov-chelovek-i-biblioteka>

⁷³ <http://lib.rus.ec>

Если автор или издатель обращаются к Мошкову с личным письмом, просят снять произведение с сайта, он снимает его немедленно. Прецедентов таких немного, гораздо чаще литераторы, наоборот, хотят публиковаться в библиотеке. За 20 с лишним лет существования «Библиотеки» все, кто хотел удалить контент, сделали это. Тем не менее, в 2004 г. к ней был подан первый иск за нарушение авторских прав. Останкинский районный суд вынес решение в пользу одного из «обиженных» авторов, признав «нарушение неимущественного права истца на обнародование произведения», но в выплате компенсации за нарушение авторского права отказал. Решение «взыскать с Мошкова Максима Евгеньевича в пользу Геворкяна Эдуарда Вачагановича в счёт компенсации морального вреда 3000 рублей» сам «ответчик» прокомментировал так: «Геворкян разрешил держать его тексты в библиотеке. Мне он писал, что никаких проблем нет, что снимать его тексты из библиотеки нет необходимости, а сам через две недели написал заявление в суд. Тексты лежали с его разрешения, а в суде он заявил, что поскольку договора на бумаге не подписано, то никакого разрешения не было, «что он там говорил, он не помнит, и вообще был пьян». Я здраво осознаю, что сотни авторов, которые разрешили мне держать их тексты в библиотеке, дали свои разрешения в устной форме при личной встрече, или по email, без подписания официальных бумажек. И любой из них, в принципе, может поступить так же, как поступил Геворкян — отречься от своего честного слова и засудить меня за то, что я когда-то ему поверил и не стал страховаться формальными бумагами. Для меня это стало уроком — я понял, что на жизненном пути встречаются люди, которым нельзя верить на слово»⁷⁴. По этому поводу можно прочесть статью А. Анненкова «30 серебряников сторублёвок Геворкяна»⁷⁵ и материал А. Экслера⁷⁶; а также мнение противоположной стороны⁷⁷.

Ситуация с авторскими правами мешает работе библиотеки. «Каждый раз, когда ты кладёшь новую книжку неизвестного тебе автора, неизвестно, как он к

⁷⁴ http://www.webplanet.ru/news/short_interview/2005/4/1/moshkoff_comment.html

⁷⁵ <http://www.babr24.com/?ADE=20846>

⁷⁶ <https://www.exler.ru/reviews/gevorkyan-moshkov-karikatura-fotografam-kalinovskij-psevdonim.htm>

⁷⁷ <https://www.km.ru/glavnoe/2005/03/30/obshchestvo/internet-pirat-moshkov-nakazan-sudom>

этому отнесётся или через сколько времени это ударит, — каждый раз это маленький стресс. Меня сейчас интересуют книжки, которые давным-давно напечатаны, а в электронной форме отсутствуют». Интерес активно пополнять библиотеку пропадает. «С авторскими правами всё очень сложно и местами абсурдно», — говорит Мошков и ссылается на историю, связанную с поэтом Николаем Гумилёвым, расстрелянным в 1921 году. Казалось бы, прошло семьдесят лет, и можно было бы оцифровать его поэтическое наследие. Но в момент расстрела он был поражён в своих правах. Права на литературные произведения отошли к государству. Его реабилитировали в 1992 году. Соответственно с этого момента ждать придется ещё семьдесят лет⁷⁸. Многие проблемы может снять использование открытых лицензий. Упомянутый закон №35-ФЗ дополнил Гражданский кодекс статьей 1286.1 «Открытая лицензия на использование произведения науки, литературы или искусства», которая вступила в силу 1 октября 2014 года. Эта важная тема заслуживает отдельного разговора.

Итак, Мошков отошёл от активного пополнения библиотеки, но поддержка её функционирования требует немалых ресурсов. «Оценочно на «Самиздате» 30 Гб литературы <это по состоянию на 2014 г. – ЮП>, этот объём прирастает по 10 Гбайт в год. Раз в несколько лет приходится с этим что-то делать. На некоторых из сопутствующих сервисов я поставил рекламу, она понемногу продаётся, обеспечивая около 40 тыс. рублей в месяц. Этих денег хватает на зарплату программиста и оплату хостинга, а также на то, чтобы раз в полгода купить новый сервер. В итоге все сервисы выходят на ноль»⁷⁹. Об этом же он говорил в интервью «Российской газете» (федеральный выпуск № 294(6865) от 27.12.2015)⁸⁰:

- На чём же зарабатывает Ваша библиотека?

- Я содержу её на свои деньги. Примерно так же, как государство содержит «бумажные» библиотеки.

- Библиотека Максима Мошкова не приносит самому Максиму Мошкову никакого дохода?

- К сожалению, нет.

⁷⁸ <https://yeltsin.ru/news/sozdatel-libru-maksim-moshkov-est-li-budushee-u-bibliotek>

⁷⁹ <http://www.unkniga.ru/innovation/tehnology/3379-maksim-moshkov-socialnye-seti-voruyut-avtorov-i-chitateley.html>

⁸⁰ <https://rg.ru/2015/12/28/moshkov.html>

- Получается, что, содержа библиотеку за свой счёт, пусть даже это недорого, Вы занимаетесь благотворительностью?

- Ну, если хотите, можно так сказать.

- Какую часть Вашего преподавательского заработка Вы тратите на Вашу библиотеку?

- Если бы я курил, я на курение тратил бы больше.

- Ну а всё же?

- Главный расход библиотеки – это зарплата программиста. Я плачу ему ежемесячно.

- Самопубликация в Вашей библиотеке – она тоже бесплатная?

- Да.

В последнее время у владельцев сайтов появилась ещё одна проблема. «Роскомнадзор выискивает везде пропаганду экстремизма, насилия, суицида, наркотиков и всего остального. Я это хорошо знаю – мне на сайт пришло за год около 20 указивок от Роскомнадзора с адресами страничек, которые надо заблокировать. Например, стишок про несчастную любовь, где была фраза «Я жить без тебя не могу». Ещё надо было удалить юмористическую заметку про «рецепт ядерной бомбы для домохозяек». У другого текста совпали слова в названии с заголовком текста из списка запрещённых, хотя содержание не совпало ... Десять лет назад «Самиздат» по довольно нелепой случайности целиком попал в список экстремистских сайтов. Выбраться было вообще невозможно, пришлось перенести сайт полностью. У тех, кто запрещает, нет ни имён, ни фамилий, ни мозгов, они не понимают, что делают. Каждый раз думаю об одном: на сайтах моих проектов есть гостевые книги, и я знаю, что готов выкинуть кого-то, забанить – и рука моя не дрогнет. Каждому автору «Самиздата» я дал механизм, позволяющий управлять тем, что ему пишут»⁸¹.

Ещё о переводе текстов в электронную форму. Оцифровать уже существующие архивы в десять раз дешевле, чем содержать бумажные библиотеки по всей стране. Но помимо проблемы с авторскими правами есть другая потенциальная

⁸¹ <https://www.colta.ru/articles/society/21190-maksim-moshkov-budet-belyy-internet-obsluzhivayuschiy-po-pasportam-i-chernyy-na-obochine>

опасность. В Библиотеке конгресса США оцифровали большой сегмент периодики. Периодику не сохранили. Но и цифровой архив сохранить не удалось, технологии шагнули вперед, и с носителями возникли проблемы. Никто не знает, что будет с цифровыми данными через сто лет⁸².

Максим Мошков протестует, когда его библиотеку называют первой и самой большой в интернете. «Было и до того много чего интересного. Но это действительно самая древняя из уцелевших библиотек. Хотя, конечно, не первая. Я стал единственным, кто не бросил это дело до сих пор. Поэтому мне достались лавры всех первых собирателей ... Она уже давно не самая большая».

Отметим: сейчас в интернете существует не менее 5000 электронных библиотек, содержащих огромный массив русскоязычных текстов, охраняемых авторским правом и попавших туда без ведома издателя и писателя⁸³.

Библиотека Мошкова много лет сотрудничает с издательствами. Ещё в 2003 году «Вагриус» издал книгу рассказов «Последняя война», написанных авторами из библиотеки⁸⁴. За ней последовали десятки других печатных изданий. Можно согласиться с утверждением: Максим Мошков, хоть и не писатель, сделал для русской литературы больше, чем дюжина иных сочинителей⁸⁵.



М. Е. Мошков (справа). Москва, апрель 2021.

⁸² <https://yeltsin.ru/news/sozdatel-libru-maksim-moshkov-est-li-budushee-u-bibliotek>

⁸³ https://www.gazeta.ru/comments/2009/09/22_x_3263625.shtml

⁸⁴ http://samlib.ru/r/romanow_w_e/pf3.shtml

⁸⁵ <https://godliteratury.ru/articles/2016/10/13/maksim-moshkov-chelovek-i-biblioteka>

СПИСОК ЛИТЕРАТУРЫ

1. Информационные ресурсы Internet. Краткий справочник // «Технологии электронных коммуникаций», т. 68. М.: Эко-Трендз, 1996.
 2. *Горбунов-Посадов М.М.* Интернет-активность как обязанность учёного. Издательские решения, 2017 63 с. ISBN 978-5-4483-7792-1
URL: <https://www.keldysh.ru/gorbunov/duty.htm>
 3. Компьютерные сети и их информационные ресурсы. Препринт ЦЭМИ РАН. М.: 1995. 72 с.
 4. *Ларин И.* Авторское право сегодня – это абсурд // «Университетская книга». 2010. № 7.
 5. Мировая сеть Internet: применение в науке и бизнесе // «Технологии электронных коммуникаций», т.59. М., Эко-Трендз, 1994.
 6. *Поляк Ю.Е.* Информационные ресурсы для популярной науки. // «Информационные ресурсы России». 2017. № 2. С. 9–12.
 7. *Поляк Ю.Е., Сигалов А.В.* Жёлтые страницы Internet'98. Русские ресурсы. СПб.: «Питер», 1998. 600 с.
 8. *Hart M.* The History and Philosophy of Project Gutenberg. August 1992 [Электронный ресурс] // Project Gutenberg.
URL: https://www.gutenberg.org/about/background/history_and_philosophy.html
 9. *Lebert M.A.* History of Project Gutenberg from 1971–2005 [Электронный ресурс] // Project Gutenberg News. 2008. February 2.
URL: <http://www.gutenbergnews.org/20080202/nef-pg-1971-2005-lebert-en/>
-

NOTES ON THE FIRST DIGITAL LIBRARIES

Y. E. Polyak^[0000-0001-8411-335X]

Central Economics and Mathematics Institute of the Russian Academy of Sciences, 47 Nakhimovski Pr. Moscow, 117418 Russia

polak@cemi.rssi.ru

Abstract

Digital libraries in the XXI century have become the most important sources of information in all branches of science and technology, the main direction of development of world librarianship, any information activities. The first electronic library was Project Gutenberg, founded by Michael Hart exactly 50 years ago. It became world famous only in the 90s, after the appearance of Internet and the first browsers. Hart dreamed of 10,000 digitized books, and by his admission, he achieved all the goals that he set for himself. In Russia, one of the first and most significant is Maxim Moshkov's library (1994). This paper traces the stages of creation of these libraries and the biographies of their creators; the difficulties in their work are analyzed. Special attention is paid to the issues of copyright duration legislation.

Keywords: *digital libraries, Project Gutenberg, Michael Hart, Moshkov Library, copyright issues.*

REFERENCES

1. Информационные ресурсы Internet. Kratkij spravochnik // «Tekhnologii elektronnyh kommunikacij», t. 68. M.: Eko-Trendz, 1996.
2. *Gorbunov-Posadov M.M.* Internet-aktivnost' kak obyazannost' uchyonogo. Izdatel'skie resheniya, 2017 63 s. ISBN 978-5-4483-7792-1
URL: <https://www.keldysh.ru/gorbunov/duty.htm>
3. Komp'yuternye seti i ih informacionnye resursy. Preprint CEMI RAN. M.: 1995. 72 s.
4. *Larin I.* Avtorskoe pravo segodnya – eto absurd // «Universitetskaya kniga». 2010. № 7.
5. Mirovaya set' Internet: primenenie v nauke i biznese // «Tekhnologii elektronnyh kommunikacij», t. 59. M., Eko-Trendz, 1994.

6. *Polyak Y.E.* Informacionnye resursy dlya populyarnoj nauki // «Informacionnye resursy Rossii». 2017. №2. S. 9–12.

7. *Polyak Y.E., Sigalov A.V.* Zhyoltye stranicy Internet'98. Russkie resursy. SPb.: «Piter», 1998. 600 s.

8. *Hart M.* The History and Philosophy of Project Gutenberg. August 1992 [Elektronnyj resurs] // Project Gutenberg.

URL: https://www.gutenberg.org/about/background/history_and_philosophy.html

10. *Lebert M.A.* History of Project Gutenberg from 1971-2005 [Elektronnyj resurs] // Project Gutenberg News. 2008. February 2.

URL: <http://www.gutenbergnews.org/20080202/nef-pg-1971-2005-lebert-en/>

СВЕДЕНИЯ ОБ АВТОРЕ



ПОЛЯК Юрий Евгеньевич – ведущий научный сотрудник Центрального экономико-математического института РАН (Москва). Подробнее: <http://computer-museum.ru/articles/sovet-muzeya/561/>

Yuri Evgenievich POLYAK – Candidate of Economic Sciences, Leading Researcher, Central Economics and Mathematics Institute. Moscow, Russia. More detailed: <http://computer-museum.ru/articles/sovet-muzeya/561/>

email: polak@cemi.rssi.ru

ORCID: 0000-0001-8411-335X

Материал поступил в редакцию 20 октября 2021 года