

СОДЕРЖАНИЕ

Введение.....	9
Глава 1. Тенденции развития и образ будущего научных коммуникаций	13
1.1. Тенденции развития научных коммуникаций.....	13
<i>Цифровизация</i>	13
<i>Открытая наука</i>	14
<i>Система оценки научных результатов</i>	15
<i>Коллаборации</i>	16
<i>Будущее научных библиотек и информационных центров</i>	17
<i>Тематические информационные центры</i>	19
<i>Публикации и данные</i>	20
<i>Перспективы развития информационных языков</i>	21
<i>Экономика научных коммуникаций</i>	21
<i>Модификация законодательства</i>	23
<i>Единое российское электронное пространство знаний (ЕРЭПЗ)</i>	23
1.2. Инициативы мировой научной общественности по развитию научных коммуникаций в условиях открытой науки	24
<i>Исходные документы по открытой науке</i>	24
<i>Открытый доступ и оценка научных исследований. Декларация DORA</i>	25
<i>Лейденский манифест</i>	28
<i>Гагская декларация по открытию знаний в цифровую эпоху</i>	32
<i>Конкордат по открытым научным данным</i>	32
<i>Амстердамский призыв</i>	34
<i>Призыв Жюсье</i>	35

<i>План S и рекомендации по его реализации</i>	35
<i>Будущее научных коммуникаций и изданий</i>	38
1.3. Выводы и рекомендации для России	44
<i>Общие выводы</i>	44
<i>Выводы в отношении категорий участников</i>	44
<i>Выводы относительно оценки научных результатов</i>	45
Глава 2. Научно-техническая информация	
в России сегодня	46
2.1. Введение	46
2.2. Краткий обзор мирового опыта	47
2.3. Общая характеристика отечественных ресурсов НТИ.....	51
2.4. Федеральные органы НТИ	53
2.5. Исследование отечественных ресурсов по управлению наукой	59
2.6. Инвентаризация владельцев ресурсов НТИ	60
2.7. Инвентаризация ресурсов	64
2.8. Выводы.....	67
Глава 3. Мониторинг информационных ресурсов	
социально-гуманитарных наук	68
3.1. Цели и задачи мониторинга информационных ресурсов.....	68
3.2. Структура научной инфосферы.....	69
3.3. Определение границ сферы мониторинга	73
3.4. Тематические классификации социальных и гуманитарных наук.....	73
3.5. Тематические границы мониторинга	77
3.6. Функциональная классификация информационных ресурсов	78
3.7. Методические вопросы инвентаризации информационных ресурсов.....	84
3.8. Сбор информации в ходе мониторинга	87
3.9. Результаты мониторинга и состав информационных ресурсов	89
<i>Архивные фонды</i>	89
<i>Фонды традиционных библиотек</i>	90
<i>Музейные фонды</i>	91
<i>Периодические, продолжающиеся и сериальные издания (ППСИ)</i>	92
<i>Вторичные ресурсы</i>	92
<i>Первичные документальные ресурсы</i>	93

<i>Информационные системы</i>	94
<i>Информация о научных событиях</i>	95
<i>Персональные ресурсы</i>	95
<i>Медиаресурсы</i>	95
<i>Прочие ресурсы Интернета</i>	96
3.10. Навигатор информационных ресурсов по общественным наукам.....	96
<i>Логическая структура НИРОН</i>	97
<i>Перспективы развития НИРОН</i>	104
3.11. О путях оптимизации академических социогуманитарных информационных ресурсов.....	105
3.12. Основные выводы по результатам мониторинга	117
Глава 4. Государственные решения в области научной информации (2018–2019)	120
4.1. О государственных решениях	120
4.2. Поручения Президента РФ	121
4.2. Законопроект «О научной и научно-технической деятельности в Российской Федерации»	121
4.3. Цифровые платформы Минобрнауки.....	127
4.4. Национальная электронная библиотека.....	132
4.5. Национальный энциклопедический портал.....	133
4.6. Другие правительственные действия в сфере научной информации	134
Глава 5. Единое российское электронное пространство знаний – постановка задачи	137
5.1. Единое российское электронное пространство знаний в «Положении о Национальной электронной библиотеке» ...	137
5.2. Некоторые замечания о понятии ЕРЭПЗ	139
5.3. Общие принципы создания ЕРЭПЗ	142
5.4. Цели и задачи ЕРЭПЗ	144
5.5. Этапы формирования ЕРЭПЗ.....	145
<i>Первый этап – мониторинг инфосферы</i>	145
<i>Второй этап – оценка и отбор научных ресурсов</i>	146
<i>Третий этап – формирование ядра ЕРЭПЗ</i>	149
<i>Четвертый этап – выявление новых знаний</i>	151
5.6. Функциональный состав участников ЕРЭПЗ	151
<i>Головная организация</i>	152
<i>Ведущие организации по отраслям науки</i>	152

<i>Операторы основных научно-образовательных электронных ресурсов</i>	153
<i>Службы формирования ядра ЕРЭПЗ и выявления нового знания</i>	154
5.7. Технологическая платформа ЕРЭПЗ	154
5.8. Организационные принципы функционирования ЕРЭПЗ	154
5.9. Экономическая модель ЕРЭПЗ	156
5.10. Правовые основы ЕРЭПЗ	156
5.11. Международное сотрудничество при создании и развитии ЕРЭПЗ	157
5.12. ГСНТИ и ЕРЭПЗ	158
Глава 6. Интеграция лингвистических информационных ресурсов	160
6.1. Введение	160
6.2. Анализ зарубежного опыта	161
<i>Общая характеристика</i>	161
<i>ELRA – Европейская ассоциация языковых ресурсов</i>	162
<i>OLAC – Сообщество открытых лингвистических архивов</i> ..	164
<i>Стандарты</i>	165
<i>Рекомендации</i>	166
<i>Методики и проекты</i>	166
<i>TELRI – Трансъевропейская инфраструктура лингвистических ресурсов</i>	166
<i>CLARIN – Общие языковые ресурсы и технологическая инфраструктура</i>	167
<i>LDC – Консорциум лингвистических данных</i>	172
<i>ISCA – Международная ассоциация речевого общения</i>	173
<i>SIL – Летний институт лингвистики</i>	174
<i>Национальные центры ЛИР</i>	174
<i>ORTOLANG – Открытые ресурсы и инструменты для обработки языка</i>	175
<i>Терминологические веб-сайты и банки данных</i>	177
<i>Каталоги ЛИР</i>	178
6.3. Концепция создания Центра лингвистических ресурсов РАН	178
<i>Постановка задачи</i>	178
<i>Определения</i>	179
<i>Состояние создания ЛИР в РАН</i>	179
<i>Типология лингвистических ресурсов</i>	181

<i>Активные формы – алгоритмы, модели, программы, базы знаний</i>	181
<i>Задачи Центра лингвистических ресурсов</i>	181
<i>Этапность создания и развития Центра лингвистических ресурсов</i>	182
<i>Целесообразность интеграции и агрегации ЛИР по отдельным типам</i>	183
<i>Корпуса</i>	184
<i>Организационные аспекты создания Центра лингвистических ресурсов</i>	186
<i>Технологические аспекты функционирования Центра лингвистических ресурсов</i>	187
<i>Экономические аспекты создания и функционирования Центра лингвистических ресурсов</i>	187
<i>Правовые аспекты создания и функционирования центра</i> ...	188
6.4. Навигатор информационных ресурсов по языкознанию (НИРЯЗ)	189
<i>Структура и функциональность НИРЯЗ</i>	189
<i>Содержание информации, включенной в НИРЯЗ</i>	189
<i>Рубрикатор НИРЯЗ</i>	191
<i>Языковой фасет НИРЯЗ</i>	192
6.5. Методика оценки качества лингвистических электронных ресурсов	193
<i>Методы оценки ресурсов</i>	193
<i>Оценка качества сайта</i>	194
<i>Параметры оценки сайтов</i>	194
<i>Характеристики оценки качества сайта</i>	196
<i>Методика оценки качества сайта</i>	201
<i>Примеры оценки сайтов</i>	204
6.6. Онтология поисковых терминов по лингвистике	207
<i>Участие в создании онтологии по комплексному гранту</i>	207
<i>Онтология поисковых терминов по лингвистике (ОПТЕЛ)</i> ..	208
Заключение	210
Литература и электронные ресурсы	213
Приложения	223
Приложение 1. Общественные науки в ГРНТИ	223
Приложение 2. Фрагмент номенклатуры ВАК по общественным наукам	223
Приложение 3. Рубрикатор РГНФ	224
Приложение 4. Рубрикатор РНФ	225

Приложение 5. Рубрикатор FOS (ОЭСР) по социальным и гуманитарным наукам	225
Приложение 6. Рубрикатор WOS по социальным наукам	226
Приложение 7. Рубрикатор WOS по искусству и гуманитарным наукам	228
Приложение 8. Фрагмент рубрикатора Scopus	228
Приложение 9. Функциональная классификация типов ресурсов по результатам мониторинга	230
Приложение 10. Упрощенная функциональная классификация, используемая в Навигаторе НИРОН	235
Приложение 11. Карта LRE [электронная карта лингвистических ресурсов].....	236
Приложение 12. Участники CLARIN.....	239
Приложение 13. Индивидуальные веб-сервисы членов CLARIN.....	240
Приложение 14. Топ-20 терминологических служб и банков данных.....	240
Приложение 15. Мировые и зарубежные каталоги ЛИР.....	242
Приложение 16. Российские каталоги ЛИР.....	243
Приложение 17. Рубрикатор по языкознанию НИРЯЗ.....	244
Приложение 18. Таблица перекодировки кодов ГРНТИ в Рубрикатор НИР.....	245
Приложение 19. Указатель языков и языковых групп в НИРЯЗ.....	246
Приложение 20. Техническое задание на формирование. Онтологии поисковых терминов по лингвистике.....	249

Введение

Исследования и разработки последних 20–25 лет в области научной коммуникации и научной информации свидетельствуют об отсутствии у российского научно-информационного сообщества общего видения будущего научной коммуникации, которое определяется тенденциями ее развития.

Под научной коммуникацией здесь понимаются все формы информационного взаимодействия в научной сфере – как традиционные (издательская, библиотечная, научно-информационная, архивная), так и современные, реализуемые в цифровой среде, прежде всего в Интернете (сайты, порталы, электронные библиотеки, базы и банки данных, репозитории, социальные сети и др.).

В мировой практике регулярно делаются попытки сформулировать коллективное мнение по образу будущего научных коммуникаций. Последним примером может служить доклад группы европейских экспертов «Будущее научных изданий и научных коммуникаций» [127], опубликованный в январе 2019 г. и подробно рассмотренный в первой главе книги.

Подчеркнем, что авторы этого документа, как и многих других коллективных концептуальных документов в области научных коммуникаций, в центр проблемы оптимизации научных коммуникаций ставят задачу адекватной оценки научных результатов. В частности, признается, что существующие методы оценки на основе импакт-факторов журналов и вообще библиометрические методы неудовлетворительны. Но изменение этих методов требует существенных социальных инноваций в области организации науки.

Попробуем сформулировать некоторые вопросы, ответы на которые должны определить образ будущего научных коммуникаций.

- Вся ли научная коммуникация через 10–15–20 лет будет цифровой?

- Является ли переход к открытой науке безусловным?

- Как будет развиваться система оценки научных исследований?
- Какую роль в научной коммуникации должны играть научные и вузовские библиотеки? Каково будущее научных библиотек?
- Нужны ли в будущем специализированные центры научной информации? Какими будут их функции? Есть ли будущее у библиографических и реферативных служб?
- Как будут модернизироваться научные издательства в условиях открытой науки и в единой цифровой среде?
- Электронные научные журналы и открытые репозитории: каково их соотношение в обозримом будущем?
- Какова перспектива существующих российских ресурсов научной информации?
- Как будет меняться соотношение централизованных (универсальных) и отраслевых (тематических) ресурсов и систем научной информации? Тот же вопрос относится к международным и национальным ресурсам и системам научной информации.
- Есть ли реальная перспектива у Единого российского электронного пространства знаний (ЕРЭПЗ)? Имеет ли смысл говорить о воссоздании Государственной системы научно-технической информации (ГСНТИ)?
- Публикации и данные: будет ли меняться их соотношение?
- Как будут развиваться информационные языки, онтологии, поисковые машины?
- Какая экономическая модель будет определять научную коммуникацию?
- Удастся ли модернизировать систему копирайта?

Безусловно, без внятного ответа на эти и многие другие вопросы усилия по созданию научно-информационных ресурсов, сервисов и институций, на которые сейчас затрачивается большой объем общественного труда, окажутся потраченными впустую.

В то же время анализ государственных решений в области научной информации показывает, что у органов государственной власти и руководства отечественной наукой отсутствует сколько-нибудь четкий план развития научной инфосферы, хотя проблемам цифровизации науки, как и экономики в целом, в правительственных документах уделяется большое внимание.

Особый интерес вызывает сформулированная в ряде нормативных документов идея создания Единого российского электронного пространства знаний. Если данная идея действительно может быть реализована, то представляется очевидным, что планы цифровизации науки и образования, как и развития научной инфраструк-

туры в целом, должны быть полностью ориентированы на достижение этой цели. Однако в реальности это не так. Хотя, по мнению многих авторитетных специалистов, эта идея не имеет шансов на реализацию при современном состоянии научно-информационного сообщества, автор считает, что обсуждение путей создания единого пространства знаний является хорошим способом организации дискуссии по сформулированным выше проблемам.

В настоящей книге делается попытка описать ситуацию в контексте поставленных вопросов и сформулировать свою позицию по вопросу создания единого электронного пространства знаний. В основном автор опирается на исследования научных коммуникаций в сфере социальных и гуманитарных наук. В частности, предлагается модель создания пространства знаний на примере лингвистики, лидирующей в России среди гуманитарных наук в области создания информационных ресурсов и систем (цифровой гуманитаристики). В то же время в российской компьютерной лингвистике особенно очевидно отсутствие координации и общего плана развития в этой сфере по сравнению с ситуацией за рубежом.

Монография имеет следующую структуру.

В первой главе рассмотрены основные тенденции развития научных коммуникаций и информационных ресурсов, отражение этих тенденций в коллективных декларациях и докладах зарубежных ученых.

Во второй главе рассмотрено общее состояние научно-технической информации в России. Описываются органы и организации, владеющие информационными ресурсами, отражающими результаты научных исследований, а также ресурсы, предназначенные для управления наукой и науковедческих исследований. Приводятся результаты каталогизации ресурсов по науковедению и управлению наукой.

В третьей главе описывается состояние информационных ресурсов академических учреждений России по социальным и гуманитарным наукам, в основном на базе мониторинга, проведенного автором в ИНИОН РАН в 2017–2018 гг.

Четвертая глава содержит обзор и анализ правительственных решений последних лет, связанных с развитием и регулированием научной инфосферы. Констатируется несогласованность этих решений.

Пятая глава посвящена центральной проблеме развития научного информационного пространства в обозримом будущем. Обсуждается возникшая в последние годы и зафиксированная в

ряде официальных документов идея создания Единого российского электронного пространства знаний.

В заключительной шестой главе представлены результаты исследования возможной модели интеграции научных информационных ресурсов на примере языкознания, выполненного в рамках проекта, поддержанного Российским фондом фундаментальных исследований.

Глава 1.

ТЕНДЕНЦИИ РАЗВИТИЯ И ОБРАЗ БУДУЩЕГО НАУЧНЫХ КОММУНИКАЦИЙ

1.1. Тенденции развития научных коммуникаций

Цифровизация

Наиболее очевидной тенденцией изменения системы научных коммуникаций является цифровизация, т.е. переход многих видов научных коммуникаций в цифровую форму.

В начале XIX в. многим специалистам казалось, что развитие информационных технологий приведет к быстрому и полному переходу научных коммуникаций в цифровую форму. Можно вспомнить обещание руководства Google, сделанное в 2005 г., оцифровать к 2020 г. *все* книги, изданные в мире. По оценке той же компании, их в мире около 150 млн. Однако к настоящему времени оцифровано не более 25 млн, а доля потребления электронных книг составляет в России 5–6% [52].

В других формах научной коммуникации – периодике, библиотечных фондах, архивах – цифровизация также не стала полной. Более подробные данные о динамике цифровизации научной коммуникации в России представлены в монографии [22].

На основе этих данных можно сделать очевидный вывод: в обозримом будущем печатная книга и журнал не исчезнут, хотя их доля будет постоянно сокращаться. Но динамика сокращения востребованности печатных форм коммуникации вряд ли превысит 5–10% в год. Некоторые виды научных документов, впрочем, уже полностью переведены в электронную форму (диссертации, отчеты). В целом рост объема цифрового контента в Интернете составил в 2017 г. 12% [113].

Таким образом, в обозримом будущем цифровая и традиционная формы научных коммуникаций будут сосуществовать, с постепенным доминированием цифровой формы.

В то же время эта очевидная тенденция совершенно недостаточно учитывается в программах и концептуальных документах, определяющих научную информационную политику, а самое главное – в представлении акторов информационной отрасли. Это напрямую затрагивает издательскую отрасль – и в части научной периодики, и в части книжной политики. Но главным образом эта тенденция меняет информационно-библиотечную и архивную сферы. Однако у научной общественности России нет общего мнения о формах и способах организации этих отраслей в новых условиях цифровой эпохи.

Открытая наука

В последние годы мировая научная общественность, особенно в Европе и в США, активно обсуждает переход к открытой науке, и в связи с этим – будущее открытых научных коммуникаций в цифровой среде. В результате появился ряд деклараций, экспертных докладов и других коллективных документов на эту тему. Особый резонанс получило решение Европейской комиссии об открытии с 2020 г. всех публикаций по результатам исследований, финансируемых европейскими правительствами и фондами [134].

Для реализации этого решения в сентябре 2018 г. были приняты так называемый план S, разработанный коалицией научных организаций 13 стран Евросоюза (коалицией S) [82], и рекомендации по его применению [128].

По вопросу о присоединении России к плану S имеются разные мнения. С одной стороны, высказывается полная поддержка этого плана [56; 59], к которой присоединяется и автор этой книги. С другой стороны, на заседании Президиума РАН в марте 2019 г., когда представитель Евросоюза призвал Россию присоединиться к плану S, вице-президент РАН А.Р. Хохлов не поддержал это предложение. Действительно, в планах руководящих органов науки России нет программы перехода к открытому доступу, хотя имеется поручение Президента РФ подготовить доклад о степени открытости научных организаций [81]. Таким образом, вопрос о динамике движения российских научных коммуникаций к открытому доступу остается нерешенным.

Система оценки научных результатов

Авторы основных коллективных документов по развитию научных коммуникаций приходят к выводу, что основная проблема их развития связана с принятыми сегодня в управлении наукой методами оценки научных результатов.

Главный вывод декларации DORA гласит:

«Не используйте такие количественные журнальные показатели, как импакт-фактор, в качестве показателя качества отдельных научных статей при оценке вклада конкретных ученых, при приеме на работу, продвижении по службе или при принятии решений о финансировании» [141].

Аналогичный вывод сделан в коллективном призыве к библиоразнообразию (призыв Жюсье):

«Система оценки научных исследований должна быть тщательно реформирована и адаптирована к практике научной коммуникации» [88].

Действительно, во многих коллективных документах [118; 127] детально проанализированы негативные последствия для науки, к которым ведет использование библиометрических показателей, таких как импакт-фактор журнала, индекс Хирша, публикационная активность и др., в качестве основного метода оценки научного результата. Впрочем, это не удивительно, поскольку, согласно известному закону Гудхарта, «когда всякий показатель становится инструментом управления, прежние эмпирические закономерности, использующие данный показатель, перестают действовать».

С другой стороны, альтернативы для библиометрии пока не выработано, хотя большинство авторов сходятся во мнении, что основой для оценки научных результатов должна стать качественная экспертиза научного результата, учитывающая цели исследования.

Поскольку, как уже сказано, система оценки научных результатов существенно влияет на всю организацию научных коммуникаций, то общую структуру коммуникаций в будущем при условии изменения этой системы оценки представить сейчас достаточно трудно. Однако можно допустить, что существующие процессы аналитической и реферативной обработки научной информации могут быть модифицированы в сторону экспертизы и создания рекомендательных сервисов.

Коллаборации

Коллаборации в науке стали одной из важнейших тенденций по развитию науки в целом и научных коммуникаций в частности. Некоторые данные по состоянию коллабораций в российской науке содержатся в работе [30].

Центром нашего интереса являются коллаборации при создании научных информационных ресурсов. Конечно, прекрасным примером коллективного создания действительно массового информационного ресурса является Википедия. Хотя Википедия и продолжает подвергаться критике за спорность некоторых данных, она является наиболее посещаемым справочным информационным ресурсом как во всем мире, так и в России. Эффективными оказались и другие научные и образовательные ресурсы, использующие технологию вики, например Летописи.ру. Описание различных подходов к коллективным формам организации научной деятельности имеется, например, в работе [42].

Хорошим примером распределенной коллективной работы над научным информационным ресурсом является участие многих российских лингвистов в разметке текстов в ходе создания Национального корпуса русского языка [72].

Фактором, активно влияющим на переход к коллективным формам научной и научно-информационной деятельности, стал, конечно, переход к разным формам открытой науки, и особенно к открытым научным данным. Приведем цитату из Конкордата открытых научных данных:

«Во многих областях, таких как кристаллография, генетика, археология и лингвистика, существует ряд превосходных примеров использования открытых данных. Это обеспечивает прогресс исследований в данных областях, повышает эффективность использования ресурсов и, следовательно, улучшает финансирование науки.

В этих областях созданы практические механизмы обеспечения открытости исследований, которые повышают культуру транспарентности, защищают от мошенничества и непредвиденных ошибок, улучшают репутацию научных учреждений

Доступ к открытым данным во многих областях стимулирует развитие новых подходов в научных исследованиях, поскольку происходит объединение данных из различных источников. Это позволяет по-новому взглянуть на междисциплинарные проблемы в самых разных областях. Во многих случаях новые подходы

обеспечиваются объединением данных ряда государственных и коммерческих организаций с данными, полученными академическими исследователями.

Открытые данные могут лежать в основе инноваций, например, когда исследователи используют данные неожиданными способами или когда компании используют данные для разработки новых продуктов. Это может привести к существенным экономическим выгодам» [146].

Обзор зарубежного опыта по созданию открытых данных на примере лингвистических ресурсов приводится в главе 6.

В то же время переход к коллаборациям во многих случаях имеет негативные последствия, а также ставит новые проблемы. В частности, усиливается острота проблем, связанных с авторскими правами при использовании чужих данных, ответственностью и правомочиями при распоряжении коллективными ресурсами, распределением затрат и т.д.

Будущее научных библиотек и информационных центров

Исходя из описанных выше тенденций, можно представить себе будущее научных библиотек¹ следующим образом.

Объемы комплектования, хранения традиционных фондов и обслуживания по ним должны медленно сокращаться, впрочем, в обозримом будущем соответствующие службы ликвидироваться не должны. Одновременно должна возрасти роль служб, которые реализуют различные пользовательские сервисы на электронных ресурсах. Это учет, каталогизация и систематизация электронных ресурсов, улучшение доступа к ним, агрегация ресурсов, рекомендательные сервисы, развитие лингвистических средств поиска, персонализация поиска, создание и поддержание специализированных электронных библиотек².

Важнейшая функция научных библиотек, которые фактически должны стать научно-информационными центрами, обслужи-

¹ Понятно, что массовые библиотеки должны модифицироваться иначе, в сторону расширения их функций как культурно-досуговых центров для населения (с основной просветительской функцией).

² Специализированные электронные библиотеки в перспективе нужны будут тогда, когда они предоставляют сервисы или фонды, дополняющие универсальные электронные библиотеки.

вающими определенную тематику, регион, корпорацию, университет или научно-образовательный центр, – это экспертно-оценочная деятельность. Она должна включать анализ электронных ресурсов, их отбор и оценку, создание рекомендательных библиографий, избирательное распространение информации, другие сервисы, в которых заинтересован пользователь данной библиотеки, включая решение правовых вопросов, связанных с доступом к электронным ресурсам.

В ряде случаев научные библиотеки могли бы взять на себя некоторые издательские функции, например печать по требованию (print-on-demand). Интеграция библиотечных и издательских функций в цифровой среде, особенно в университетах, в целом представляется весьма перспективным направлением развития. Примером российской научной библиотеки, которая развивается в таком направлении, объединяя функции научно-информационного центра, вузовской библиотеки, а также многие другие, является Центральная научная медицинская библиотека Первого МГМУ им. И.М. Сеченова [110].

Библиографические продукты, как представляется автору, могут быть представлены в будущей цифровой среде в виде сводной базы данных национальной библиографии, вероятно, на основе баз данных Книжной палаты и библиографических агрегаторов, таких как ЛИБНЕТ, ЭКБСОН или АРБИКОН, а также каталога Национальной электронной библиотеки. Эта сводная база данных должна выполнять следующие функции:

- обеспечивать корпоративную каталогизацию и развитие электронных каталогов библиотек;
- обеспечивать навигационные функции – поиск местонахождения изданий;
- быть справочным аппаратом для национальной программы оцифровки;
- способствовать обеспечению сохранности книжного фонда и созданию Страхового книжного фонда.

Реферативные службы могут иметь будущее, только если они возьмут на себя функции оценки и экспертизы реферируемых публикаций, включая функции выявления плагиата. Иначе говоря, реферативные службы разумно переориентировать на обработку и оценку отечественной литературы. Такие функции наиболее естественно реализовывать в рамках репозитория научных публикаций.

Более подробное описание современного состояния российской научной библиотечно-информационной сферы представлено в главе 2.

Центральной задачей управления библиотечно-информационной сферой в России должна стать выработка оптимального соотношения между универсальными и частными (специализированными) ресурсами и сервисами научной информации, равно доступными в единой цифровой среде. Так, например, в случае электронных библиотек мы наблюдаем значительное дублирование при их создании, что означает неэффективное использование довольно ограниченных общественных ресурсов

Тематические информационные центры

Наряду с библиотеками будущего, превратившимися в научно-информационные центры, хорошие перспективы должны иметь специализированные структуры, обеспечивающие реализацию различных сервисов для определенной категории научно-информационных ресурсов.

Так, в настоящее время в рамках реализации исследовательского проекта РФФИ № 18-00-002/18 «Интеграция научно-информационных ресурсов учреждений РАН (на примере языкознания) как части единого цифрового пространства РАН» автор принимает участие в разработке концепции создания Центра лингвистических ресурсов РАН. Этот центр предполагается создать на базе служб поддержки Национального корпуса русского языка при Институте русского языка РАН. Более подробное описание принципов создания такого центра приводится в главе 6 настоящей монографии.

Возможно, в рамках социальных и гуманитарных наук должно возникнуть еще несколько тематических информационных центров (например, по социологическим или этнографическим данным, синологии, монголоведению и др.). Информационный ресурс по отечественной истории фактически уже создан на базе Президентской библиотеки им. Б.Н. Ельцина, хотя его функции в общей системе научной информации по социальным и гуманитарным наукам еще придется определить.

Публикации и данные

Основной массив научной информации, отражающей результаты научных исследований, представляет собой научные публикации, а также некоторые категории непубликуемых документов (отчеты, диссертации). Однако в современной научной коммуникации, особенно в условиях открытой науки, возрастает значение разнообразных результатов научных исследований, не прошедших традиционный путь оформления через публикации, отчеты и диссертации. Чаще всего такие результаты изначально появляются в цифровой форме (born digital). Эти результаты обобщенно стали называть «данными».

Так, в цитированном выше коллективном документе по управлению открытыми научными данными (Конкордате), который разработан консорциумом университетов Великобритании, следующим образом определяется понятие научных данных:

«Научные данные – это: статистические данные, подборки цифровых изображений, аудиозаписи, стенограммы интервью, данные обследования и наблюдения на местах с соответствующими аннотациями, интерпретации, произведения искусства, архивы, найденные предметы, опубликованные тексты или рукописи и др.» [146].

В Конкордате можно найти множество ссылок, описывающих опыт работы с научными данными.

Понятие научных данных дополняет понятие научных публикаций, которые ранее рассматривались как основная форма научной коммуникации.

Отметим, что работа с данными никогда не входила в сферу библиотечной деятельности, ориентированной на работу с публикациями, но входила в сферу интересов органов научной информации¹.

Приведем некоторые принципы управления научными данными, сформулированные в Конкордате по открытым научным данным:

- хорошее управление данными имеет фундаментальное значение для всех этапов исследовательского процесса и должно быть обеспечено с самого начала;

¹ Опыт работы с данными (с фактографической информацией) в институте информации оборонного комплекса освещен в монографии [49].

- управление данными имеет жизненно важное значение для того, чтобы данные были полезными для других, а также для сохранения данных в долгосрочной перспективе;

- данные, на которые ссылается публикация, должны быть доступны к дате публикации и должны быть представлены в цитируемой форме;

- обучение работе с данными является сферой ответственности всех заинтересованных сторон; администрирование, архивирование, манипулирование и анализ данных требуют набора навыков, отличных от тех, которые используются для сбора, генерации или измерения данных.

Перспективы развития информационных языков

Параллельно с использованием традиционных информационных языков в последние годы создавались новые современные семантические и лексикографические ресурсы, в том числе в форме тезаурусов и онтологий, словарей на основе корпусов и др. Таких ресурсов только в РАН создано уже много десятков. Их описание содержится в Навигаторе информационных ресурсов по языкознанию, описанном в главе 6, а каталоги зарубежных и отечественных лингвистических ресурсов представлены в приложениях 14–16.

Задача состоит в организации сотрудничества для разработки и развития лингвистических средств как для научных целей, так и для прикладных задач, связанных с применением семантических технологий в научной коммуникации.

Экономика научных коммуникаций

Основное содержание социально-экономической составляющей научно-информационной деятельности сегодня – это конкуренция трех способов народно-хозяйственной деятельности: государственного, рыночного (коммерческого) и волонтерского (общественного). При этом и российский, и зарубежный опыт с очевидностью показывает, что доминирование любого из секторов приводит к негативным последствиям. Поэтому задача заключается в выстраивании баланса между экономическими моделями научных коммуникаций.

Однако механизм достижения этого баланса совершенно не ясен, тем более что ведется активная борьба владельцев научных ресурсов и средств коммуникаций за доходы в рыночных условиях. Владельцы научной информации, будь то научные организации, журналы или архивы, сплошь и рядом пытаются извлечь прямую финансовую выгоду из факта владения научной информацией, независимо от того, насколько это соответствует или противоречит общественным интересам.

Проблемы оптимизации экономических моделей научных коммуникаций составляют основное содержание дискуссий, связанных с переходом к открытой науке. Достаточно подробный анализ этих проблем содержится в докладе экспертов Европейского союза о будущем научных коммуникаций [127].

Специфика России заключается в том, что у нас вообще нормативно не определена роль основных акторов научной коммуникации, к которым относятся:

- государство в лице органов власти и управления;
- Российская академия наук;
- научные фонды;
- университеты и научно-исследовательские институты;
- издательства и редакции;
- коммерческие информационные сервисы;
- общественные информационные сервисы;
- библиотеки (традиционные и цифровые);
- архивы (традиционные и цифровые);
- авторы;
- потребители научной информации.

Особенно это удивительно на фоне грандиозных планов развития цифровой экономики. Признавая теоретически неизбежный переход к экономике нового типа, в том числе к экономике знания и экономике внимания, признавая также, что научная информация относится к общественным благам, руководство российской науки и образования тем не менее до сих пор не выработало адекватный взгляд на экономику научных коммуникаций.

Вместе с тем теоретических исследований этих проблем в России было проведено немало. Упомянем, в частности, монографии А.Б. Долгина [40; 41], а также фундаментальный отчет А.Н. Козырева и его коллег [94], в которых подробно рассматривается возможность применения различных экономических моделей к организации научной коммуникации.

Однако, насколько известно автору, результаты этих исследований не нашли отражения в экономической практике научных коммуникаций. Поэтому прогнозировать развитие научной коммуникации, не имея ясного представления о ее экономике, достаточно затруднительно. Впрочем, стратегия, направленная на огосударствление многих функций в общественной сфере, реализуемая в последние годы, дает некоторые возможности для прогноза, хотя ожидаемый результат не представляется оптимальным.

Модификация законодательства

Одним из важных препятствий для развития научных коммуникаций в цифровой среде является действующее законодательство в области авторского права. Правовые проблемы развития научной инфосферы достаточно подробно рассмотрены в цитированной выше монографии «Инфосфера общественных наук России» [22]. Там же изложены предложения по изменению российского законодательства в области авторского права, которое давно не соответствует современному уровню коммуникаций. Добавим, что законодательство многих стран, особенно европейских, существенно модифицировалось в направлении обеспечения свободы распространения научного знания.

Однако, несмотря на назревшую необходимость, в России пока не торопятся вносить изменения в законодательство, хотя законодательных актов, регулирующих деятельность в Интернете, принимается более чем достаточно. Остается надеяться, что стимулом для внесения в законодательство в области авторского права изменений, необходимых для развития научных коммуникаций, станет правительственная программа «Цифровая экономика», реализация которой потребует серьезной модернизации законодательной базы.

Единое российское электронное пространство знаний (ЕРЭПЗ)

Среди предложений, которые должны формировать образ будущего, особое место занимает идея Единого российского электронного пространства знаний (ЕРЭПЗ). Эта идея провозглашена в ряде нормативных документов и представлена как перспектива

развития Национальной электронной библиотеки. Проблемам создания ЕРЭПЗ и его соотношения с системой научной информации в России полностью посвящена глава 5 данной книги.

1.2. Инициативы мировой научной общественности по развитию научных коммуникаций в условиях открытой науки

В последние годы мировая научная общественность, особенно в странах Европейского союза, активно обсуждает проблемы перехода к открытой науке и в связи с этим – будущее научных коммуникаций в цифровой среде. В результате появился ряд деклараций, экспертных докладов и других коллективных документов на эту тему. Особый резонанс получило решение Агентства по науке Европейской комиссии об открытии с 2020 г. всех публикаций по результатам исследований, финансируемых европейскими правительствами и фондами [134].

В настоящем разделе представлен обзор основных коллективных документов, освещающих проблемы перехода к открытой науке и тесно связанную с этим переходом проблему оценки результатов научной деятельности. В этих документах возникает образ будущего научных коммуникаций в условиях их функционирования в открытой цифровой среде. Главная проблема, конечно, заключается в возможности применения в российских условиях идей и принципов, провозглашаемых в коллективных документах по открытой науке.

Исходные документы по открытой науке

Первое важное международное упоминание об открытом доступе содержится в материалах Будапештской конференции по открытому доступу, проходившей в феврале 2002 г. [38].

Вскоре за ней последовала Берлинская декларация об открытом доступе к научным и гуманитарным знаниям в октябре 2003 г. [29].

Далее последовало принятие ряда документов отдельными научными союзами и организациями, среди которых нужно отметить документы Организации экономического развития и сотрудничества (ОЭСР) и Фонда открытого знания. В 2004 г. ОЭСР (Organisation for Economic Cooperation and Development, OECD)

было подписано коммюнике, в котором сформулировано требование доступности для граждан архивов всех исследований, финансируемых за счет государственных средств [142].

В 2007 г. были опубликованы «Принципы и положения для обеспечения доступа к научным данным исследований, финансируемых государством» [132]. В них говорится о том, что в связи с высокой скоростью развития Интернета и технологий, связанных с передачей, хранением и обработкой информации, появились новые возможности по совместному использованию научных данных мировым сообществом. В качестве наглядного примера в них приводится международный проект «Геном человека».

В июле 2009 г. был опубликован первый вариант *Пантонских принципов*, которые представляют собой набор рекомендаций по опубликованию научных данных, доступных для их последующего использования всеми заинтересованными лицами. В 2010 г. Пантонские принципы были уточнены и доработаны при участии членов рабочей группы «Открытые данные в науке» Фонда открытого знания [138].

В Пантонских принципах, в частности, говорится о том, что научные данные должны стать частью общественного достояния и публиковаться с соответствующей лицензией таким образом, чтобы предоставлять возможность их повторного использования.

В 2009 г. Европейская организация по ядерным исследованиям (CERN) опубликовала книжный каталог открытых данных в рамках проекта «Открытая библиотека» [136]. ЦЕРН поддерживает инициативу открытого доступа, поэтому неудивительно, что она заключила соглашение с рядом организаций, по которому научные данные, полученные при помощи Большого адронного коллайдера, публикуются в открытом доступе под лицензией Creative Commons.

Открытый доступ и оценка научных исследований. Декларация DORA

Достаточно быстро научному сообществу стало очевидно, что проблема открытого доступа теснейшим образом связана с проблемой оценки результатов научной деятельности. В 2012 г. была опубликована *Сан-Францисская Декларация об оценке научного исследования*, получившая известность под акронимом DORA [141]. С этого момента все дальнейшие документы по открытому доступу ссылаются на принципы DORA и выражают им поддержку.