

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ВСЕРОССИЙСКИЙ ИНСТИТУТ НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ

РОССИЙСКОЙ АКАДЕМИИ НАУК

(ВИНИТИ РАН)

НАУЧНО · ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 1. ОРГАНИЗАЦИЯ И МЕТОДИКА
ИНФОРМАЦИОННОЙ РАБОТЫ

ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 1

Москва 2024

ОБЩИЙ РАЗДЕЛ

УДК 001.1:[159.95+004.8]

Н.В. Максимов

Знания и информация в контексте естественного и искусственного интеллекта*

Рассматриваются обобщенный состав и структура рациональной деятельности. Модель мыслительных процессов представляется в виде взаимодействующих функциональных блоков переработки и хранения информационных структур. В контексте деятельностного подхода рассматриваются основные понятия естественного и искусственного интеллекта, а также определены характеристические свойства сущности "знания": систематизированность, предметность, квалифицированность, распределенность, гранулярность. Обосновывается необходимость понимания интеллекта в синтезе со знанием и деятельностью.

Ключевые слова: информация, знания, данные, деятельность, информационная схема познания, искусственный интеллект

DOI: 10.36535/0548-0019-2024-01-1

* Работа выполнена при поддержке Министерства науки и высшего образования РФ (Проект государственного задания № FSWU-2023-0031).

ВВЕДЕНИЕ

Сегодня искусственный интеллект (ИИ) – это приоритетное направление в информационных технологиях. Технологии искусственного интеллекта развиваются интенсивно и по многим областям дают впечатляющие успешные практические результаты. Сложились и по-разному обсуждаются и формируются понятия и направления сильного и слабого, общего и узкого интеллекта. Однако, несмотря на то, что интеллект является инструментом познания, функция понимания/познания как таковая в ИИ не определена.

Можно сказать, что глобальное назначение (сверхзадача) искусственного интеллекта – это синтез нового знания на массиве старого и развитие самого ИИ. Но в идеале цель познания (в первую очередь научного) – это открытие нового явления или его предсказание, создание принципиально новых решений, и это, так или иначе, практика, т.е. *физический* выход за пределы знания как артефакта.

Искусственный интеллект не обязательно должен быть человекоподобным. Но может ли он быть не антропоцентрическим? Развитие естественного интеллекта (ЕИ) определяется экзистенциальной целью человека – выживание во взаимодействиях с окружающей средой, в частности, путем развития самого человека и/или изменением окружающей среды. Соответственно, функции интеллекта «объединяются» в общем процессе «жизнь». При этом и непосредственно в процессах взаимодействий возникают проблемные ситуации, оперативно и итеративно выявляются цели, формулируются и реформулируются задачи, находятся альтернативные решения. В этом смысле непонятно, в чем может состоять глобальная цель ИИ. Тем не менее пока очевидна неизбежная ориентация на цели человека/общества, что позволит целенаправленно и достаточно эффективно объединить в рамках конкретных целей функции разных по своему назначению систем искусственного интеллекта.

Сознание, как и информация, сущность преходящая, существующая, скорее, как действие, порождающее результаты-макрообъекты (артефакты, изменения в окружающей среде, знания), по которым мы и делаем заключение о его наличии. По приведенному в [1] образному сравнению Джозефа Богена, сознание как ветер: увидеть и поймать его нельзя, но очевидны результаты его деятельности – гнувшиеся деревья, волны или даже цунами.

Но говоря об искусственном интеллекте, необходимо понимать, что мы рассматриваем его как макрообъект – техническую систему, компоненты которой в совокупности имеют «...способность, подобно человеку, мыслить, взаимодействовать, адаптироваться к изменившимся условиям и решать другие задачи в области обработки информации, ассоциирующиеся с естественным интеллектом человека»¹ [2].

В настоящее время основные усилия в области искусственного интеллекта сосредоточены на моделировании и разработке инструментов построения вывода,

причем преимущественно в предположении, что знания, на которых строится вывод, – это предварительно (и в известной степени субъективно) размеченный массив, образ которого локализован в самом ИИ.

Но при этом очевидно, что основная цель и назначение направления искусственного интеллекта состоит не в разработке тех или иных моделей и средств, а в достижении с их применением определенных результатов в деятельности человека. Причем изначально принимается допущение, что это целенаправленная и разумная деятельность, а это, безусловно, предполагает использование знаний и, соответственно, их воспроизведение. И в этом смысле интеллект (в том числе и искусственный) и знания надо рассматривать и применять совместно, так как они представляют собой целостный инструмент синтеза моделей/оценок, которые, в свою очередь, есть собственно информационные компоненты процессов взаимодействия (с управлением) с окружающей средой, в том числе для получения практического полезного результата в материальной сфере.

И необходимо помнить, что знания по своей природе всегда относительны и неполны. Выводимые знания должны проверяться теорией и практикой, а также применяться и передаваться другим субъектам познания для развития или пересмотра. Надо определять, что же передавать помимо полученных результатов (представленных в итоге кодом) – какой формировать сопутствующий контекст (метазнания, метаданные), чтобы данные воспринимались как действенная информация, на которой другой субъект деятельности, в свою очередь, построит новые знания. Для этого надо знать (или хотя бы представлять, моделировать), как и зачем обрабатывается поступающая информация (точнее, информационные макрообъекты – далее ИнфОб) и какие структуры, с каким содержанием и как для этого используются. Это позволитrationально решать задачи, организовывать воспроизведение и передачу знаний, а также не только развертывать на практике координацию работы агентов ИИ при решении конкретных задач, но и согласованно вести разработку искусственного интеллекта.

В [3] была предложена функционально-информационная схема когнитивных процессов, которая рассматривалась с технологической точки зрения.

В настоящей статье в рамках деятельностиного подхода мы рассмотрим свойства таких основных информационных объектов этого процесса, как «потребность», «мотив», а также «данные», «информация» и «знания».

ИНТЕЛЛЕКТ В КОНТЕКСТЕ ДЕЯТЕЛЬНОСТИ²

В [4] определяются следующие критерии ментального процесса.

1. *Разум есть совокупность взаимодействующих частей или компонентов:* разум не является неразделенной, единой субстанцией. Разум состоит из эмерджентных саморегулирующихся и самоорганизующихся частей, в приближении к которым мы не обнаруживаем ничего, что можно назвать «разум-

¹ Но, отметим, человеку не дано знать истину – у него есть только возможность приблизиться к ней. При этом человеку свойственно ошибаться. И в этом случае, одним из естественных выходов является его смерть, что «освободит» путь другим решениям и направлениям. Хотя не надо забывать, что и ошибочность – это оценка, имеющая относительную природу.

² Здесь представлена авторская точка зрения, основные положения которой сформулированы в результате анализа и обобщения многочисленных материалов публикаций по вопросам сознания и познания, обзор которых был бы для настоящей публикации неприемлемо объемным.

ным», но, тем не менее, можем признать, насколько части находятся друг с другом в *отношениях*.

2. *Коммуникация не происходит «в вакууме*, но *содержится в контексте*, который образован *презумпцией и вопрошанием*. Актуальная информация (т.е. единица информации) – это «различие, которое производит различие», «действенное различие».

3. *Ментальный процесс требует сопутствующей энергии*: ментальный процесс должен быть открытой системой по отношению к «энергии», иначе он неизменно завершится увеличением энтропии. Этот критерий является следствием критерия 2: взаимодействие по триггерному принципу инициируется информацией, т.е. различиями.

4. *Ментальный процесс требует циркулярных – или более сложных – детерминирующих (каузальных) цепей*.

5. В ходе ментального процесса результаты *воздействия различий* следует рассматривать как *трансформанты* (т.е. закодированные версии) предшествующих событий: вся действенная (эффективная) информация должна быть *сохранена*. Этот критерий признает, что обратная связь неизбежно кодируется, что согласуется с тем, как познание заключает в себе ментальное отображение внутреннего мира. Таким образом, критерий постулирует структуру «смысла» как процесса, в котором информация кодируется в новую форму и только затем становится доступной другим уровням системной организации.

6. *Описание и классификация процессов трансформации вскрывают иерархию логических типов, внут-*

ренне присущую феномену: нарушения этой иерархии инициируются элементарными типами логического вывода в системе категорий.

Таким образом, интеллект, как инструмент выполнения интеллектуальных операций, – это комплекс имеющих информационную природу³ взаимосвязанных систем, воспринимающих, обрабатывающих, хранящих и синтезирующих объекты-образы. Соответственно, интеллект надо рассматривать не только как «мозг» (*процессор+память*), но и как его состояние (знания), а также во взаимосвязи с сенсорами/эффекторами, осуществляющими связь с окружающей средой (ощущения/воздействия), в том числе использование/передачу внешних, обобществленных знаний. Более того, рассматривать «инструмент» вне контекста его применения (а также появления и развития) малопродуктивно. Это было бы подобно разработке ЭВМ без устройств ввода/вывода и безотносительно специфики ее применения: машина будет, но ее эффективность может оказаться ограниченной классом задачи или быть вовсе неприемлемой.

Приведенная на рис. 1 функционально-информационная схема когнитивных процессов⁴ составляющие основу интеллектуальной деятельности, показывает, что уже этапы, относящиеся к восприятию полученного стимула (определение и оценка значения, узнавание), зависят не только от самого стимула, но и от других сопутствующих (ассоциированных со стимулом) факторов: свойств потребности, мотивов, понимания субъектом предметной области (ПрО) и т.п.

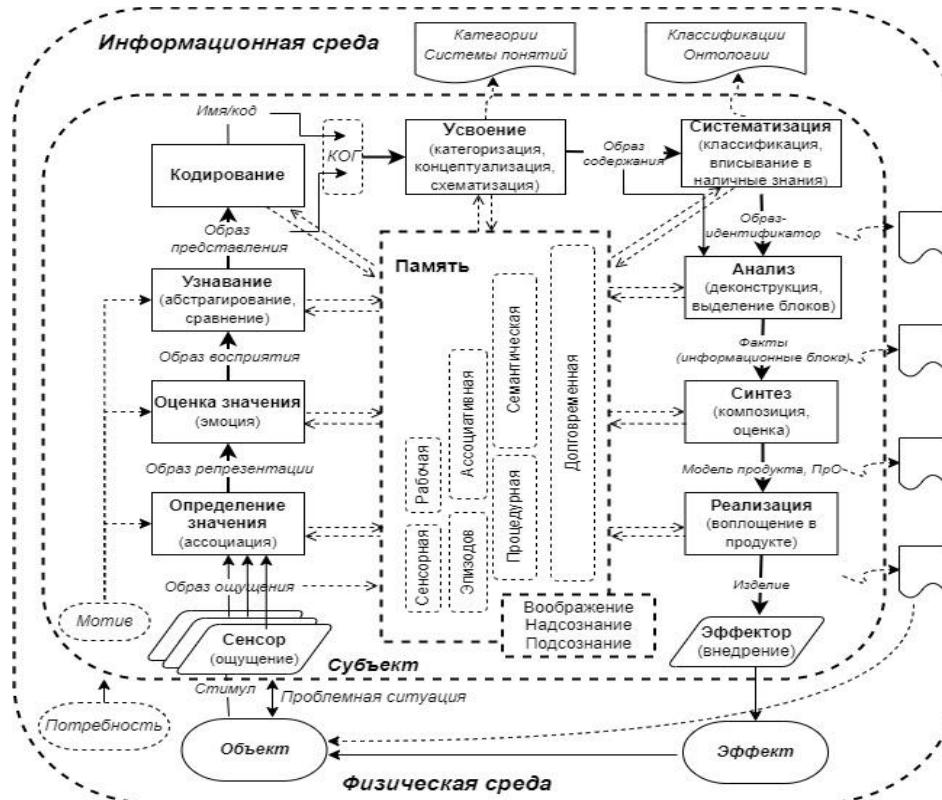


Рис. 1. Функционально-информационная схема когнитивных процессов (приводится по [3])

³ Информационная природа сознания рассматривалась в [3]. На это также, так или иначе, указывают в своих работах классики как нейрофизиологии (П.К. Анохин, Н.П. Бехтерева, С. Пинкер, Ф. Варела, Т.П. Зинченко и др.), так и кибернетики (У.Р. Эшби, Н. Винер, М. Арбид, Р. Пенроуз, Н.Н. Моисеев, Д.И. Дубровский, Д.С. Чернавский и др.).

⁴ В целом компоненты этой схемы обобщенно соответствуют компонентам схем когнитивной архитектуры ИИ, см., например, [5].

Здесь (см. рис. 1) **объект** – это воспринимаемая как целое часть реальности, используемая в конкретном (целевом) процессе и доступная для наблюдения и взаимодействия.

Потребность – это материальная или идеальная сущность, то, что имеет конечную ценность для субъекта и достижение чего выступает смыслом деятельности. Потребность – это, по сути, состояние организма, когда ему объективно требуется нечто (недостаток или излишек чего-либо, несоответствие, которое создает проблемную ситуацию).

Потребности, как нечто нужное, с функционально-системной точки зрения можно подразделить на потребности «рабочего тела» – то, что обеспечивает существование и реализацию способностей интеллекта, и потребности, связанные с развитием организма и его будущими состояниями (т.е. и рабочего тела, и интеллекта, и, возможно, ближнего окружения, с которым осуществляется взаимодействие). Именно к последним относятся характерные для млекопитающих две, по терминологии [6], «бескорыстные» потребности. Это потребность познания/любопытства и потребность сопереживания (эмоционального резонанса для эмоционального состояния другой особи). Но именно они и обеспечивают будущее устойчивое существование и возможность ускоренного развития организма и, так или иначе, социума и окружающей среды.

Потребность конкретизируется через опредмечивание: она становится *определенной* потребностью – сущностью, обладающей свойствами.

Потребность, так или иначе, ассоциируется с путями достижения целевого состояния, а также с эффектом от произошедших изменений. Такие ассоциации представляют собой **мотив** – то зачем и/или как будет осуществляться деятельность. Субъектом мотив обнаруживается посредством переживаний, характеризующихся положительными либо отрицательными эмоциями, которые не только стимулируют процесс познания, но и привносят в мышление алогичность.

Стимул – это ситуативно возникающий в процессе деятельности материальный или идеальный объект (состояние), некоторым образом ассоциирующийся с потребностью, например, предположительно входящий в структуру как часть или обладающий свойствами, приводящим к удовлетворению потребности.

В познании стимулом деятельности является неопределенность, выступающая как свойство, характеризующее проблемную ситуацию [7]. Состояние неопределенности (соотношение стимула и внутреннего знания) – это симптом возможного противоречия. Если это противоречие будет исследоваться, то оно формализуется в виде *проблемы* – состоявшегося, признанного противоречия, которое еще не имеет решения или может иметь неоднозначные решения. Такое решение может быть найдено как рациональным способом – сведением к *решаемым задачам* (методам в соответствующем классе задач), так и иррациональным – выходом за пределы существующего знания. Решение задачи тесно связано с выбором *критериев оценки* (в том числе и альтернатив, представляющих различные точки зрения и основания), например, таких, как простота, обоснование правдоподобности, синтаксические, семантические и другие.

Достижение (или не достижение) цели, т.е. полное или частичное удовлетворение потребности вызыва-

ет у человека **эмоции**, что приводит к изменениям состояния его нервной системы и таким образом к формированию оценки стимула, а, возможно, и самой потребности. Эмоция – это оценка изменения *состояния организма* в категориях системы внутренних ценностей вследствие действия стимула. Причем, на эмоциональную оценку влияет не только значимость полученного результата, но и эффективность, и «удобство» его получения, ожидаемое и актуальное воздействие на развитие субъекта, а также соответствие наличным знаниям (в том числе и реликтовым) и т.д. – все, что явно или неявно оказывает влияние на психику. Эмоция, несмотря на то, что имеет мультиodalную природу, – это целостная, не сводимая к отдельным показателям оценка.

При этом уже на постановочном этапе процесса целенаправленной деятельности (и, в частности, решения проблемы/задачи и синтеза знаний, схема которого приведена на рис. 2), на каждом шаге накапливаются ассоциированные факты – информация, представляющая собой своеобразный «питательный бульон». Эти факты, скорее всего, не использовались в изначальных формулировках проблемы/задачи/критерия, однако они могут быть полезны как альтернативные или дополнительные при оценке промежуточных и итогового решений, обеспечивая обозримость при доказательстве состоятельности решения или при реформулировании задачи, а также при достраивании и перестраивании собственных знаний. Они также формируют «контекстное поле» – некоторую содержательную основу для организации последующего применения решения и его развития

Более того, получение субъектом деятельности информации о своем состоянии дает возможность осуществлять соотнесение своего развития с развитием *целого*, а не отдельных частей, и позволяет производить уже и самооценку. В случае же ИИ – это преимущественно только построение *выдачи* по лаконичной спецификации задачи/запроса на основе накопленного обобществленного знания, а результаты вывода оцениваются вовне.

Выбор обеспечивающей управление взаимодействием с окружающей средой траектории развития системы (правильнее было бы сказать – «организма» по А. Богданову [8] основывается на использовании ценностей (критериев, смыслов, значений). А для саморазвивающихся систем, реализующих и управление собственной структурой, – это еще и решение обычно неявно сформулированной основной, экзистенциальной задачи обеспечения *сохранности организаций* системы. Отсюда следует, что наряду с системой ценностей, определяющей извне целесообразность/эффективность функционирования, должна существовать и *внутренняя* система ценностей, обеспечивающая *автономный* контроль за восприятием, обучением, выводом, а также за собственной реорганизацией.

⁵ Можно предположить, что множество цепочек ассоциированных фактов, относящихся к различным модальностям, стратам и средам познания/деятельности организма, и является основой (формирует «критическую массу») для озарений, которые характеризуются как инсайты, поскольку этому не предшествует в явной форме последовательность рассуждений, образующих вывод. Озарение – это процесс самоорганизованной критичности (по терминологии Г. Бейтсона [4]).

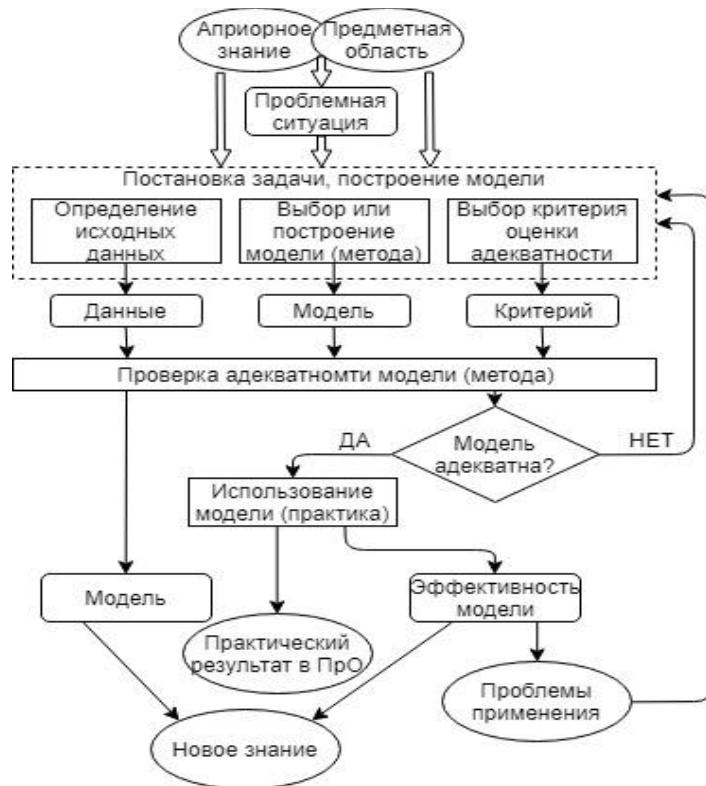


Рис. 2. Схема (информационная модель) процесса решения проблем – синтеза знаний

С технологической точки зрения познание, обучение, решение задач и проблем – т.е. то, что относится к разумной (интеллектуальной) деятельности⁶, – это, по существу, реализация циклопричинной цепи *взаимодействие со средой – интеллект – знания*. Здесь действие создает условия для связи с внешней, в итоге, с физической средой. Знания обеспечивают связь во времени и пространстве с информационной средой анализа/синтеза информации/знаний. Интеллект поддерживает «функциональную» связь знаний и деятельности в контексте ее цели, формируя решения и оценки их эффективности.

Характерная черта процесса распределенной разумной деятельности, обобщенная схема которого представлена на рис. 3, заключается в том, что получение практического результата осуществляется в соответствии с некоторой моделью. Такой процесс можно представить как предиктивный и итеративный, использующий обратные связи. Цели опреде-

ляются текущим и ожидаемым состоянием процесса (в пределе – «завершающейся технологией», приводящей цепочку стратегических и технологических решений к достижению глобальной цели); пути достижения цели – наличными знанием/опытом и ресурсами, доступными субъекту, а критерий достижения цели (и в той или иной степени сама цель) – в том числе и волей, которая определяет, насколько настойчиво будет действовать субъект, особенно в случае получения не вполне удовлетворительного результата. Существенно также и то, что получаемые промежуточные результаты или изменившиеся обстоятельства могут радикально изменить ход процесса, переопределить цели и парадигмы.

Сущность «Субъект деятельности» (см. рис. 3) функционально соответствует понятию «организм» – система систем: симбиоз разнородных систем, имеющих общую цель (точнее подсистем, относящихся к ее достижению) и некоторую общую элементную базу.

Модель – это образ «Практики» (преобразования, взаимодействия, объекта и т.п.) – то, что позволяет прогнозировать свойства и/или поведение моделируемого объекта, т.е. «Практики». Это функционально тождественное преобразование образа исходного состояния моделируемого объекта в образ целевого состояния, которое рассматривается как предопределенное прогнозируемое состояние.

Важно, что с точки зрения формирования знаний как объекта, эмоция является одной из составляющих контекста соответствующего фрагмента знаний, которая способствует их выявлению, идентификации и использованию.

⁶ Разумная деятельность инициируется и контролируется разумом, причем вначале формируется суждение, потом выполняется действие, затем производится оценка действия и выносятся суждения. Для разумной (или рациональной, рассудочной, интеллектуальной) деятельности, помимо многошаговости, характерны многоуровневость, а также интерактивность и итеративность. Реализация и контроль деятельности основываются на предопределяющей ее модели и использовании обратной связи. Деятельность строится на согласованном совмещении (по свойствам) своих основных сущностей: потребностей и мотивов субъекта, целей и средств, методов и критерии, результатов и эффектов.

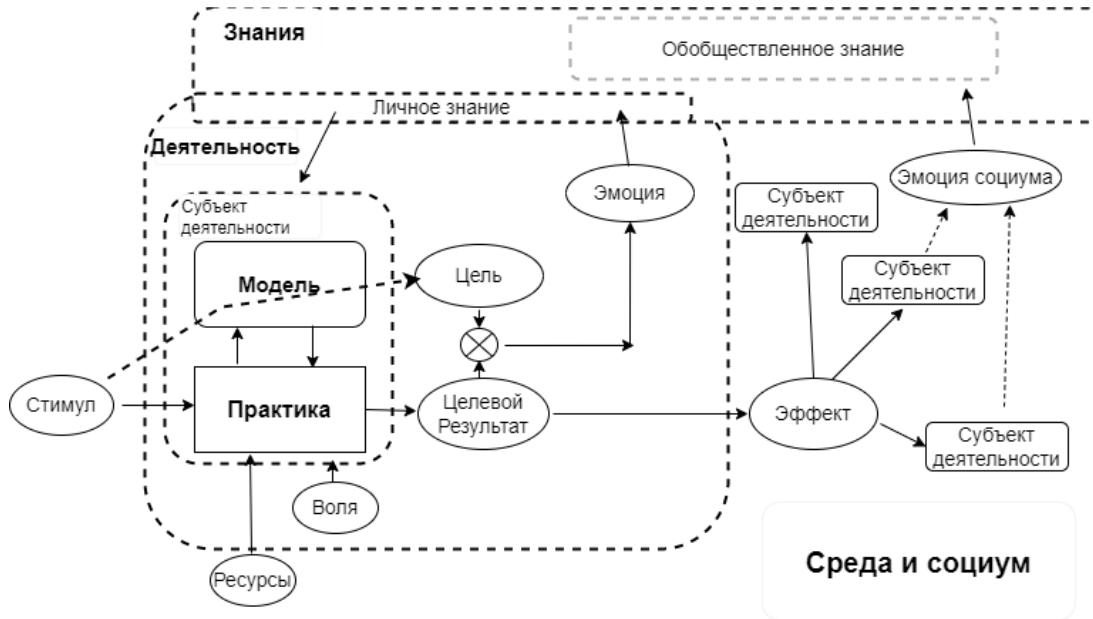


Рис. 3. Схема процесса распределенной разумной деятельности

Отметим, что построение модели определяется не только свойствами моделируемого объекта и его взаимодействий с окружающей средой, но и возможностями практической/экспериментальной проверки адекватности этой модели, включая предполагаемые необходимые измерения исходного и итогового состояния, или согласованностью (или объяснением) с ранее полученными данными.

ЗНАНИЯ

Результатом деятельности является не только собственно целевой результат – изменение состояния среды или самого субъекта, но и знания. Причем получаемое знание выступает не только как отдельный «продукт», используемый для применения или для синтеза нового знания, но и как ресурс для воспроизведения (в том числе обучения, образования) самого познающего субъекта, и, в частности, для обучения систем ИИ. Отметим, что побочным «технологическим» продуктом целенаправленной познавательной деятельности является создание коммуникационного сопряжения (языков, систем величин, категорий и понятий, а также лингвистического обеспечения), реализующего связь познаваемого и познающего.

Знания – это представление результатов интеллектуальной и практической деятельности в адекватной для восприятия форме, обеспечивающей устойчивое целевое использование и воспроизведение знаний. Роль знания – «рабочий» объект, ситуативно соответствующий субъекту и объекту деятельности. При этом декларативные знания (и лингвистическое обеспечение, в частности) выполняют роль общеприменимых «опорных точек» смысла, обеспечивающих общность представления создаваемого знания в распределенной среде и относительно которых можно строить объективно сопоставимые оценки.

Знания, даже как самоцель процесса познания, не являются продуктом только творчества: взаимодействуя с действительностью, проводя экспериментальные исследования, субъект стремится организовать их таким образом, чтобы иметь возможность «вступить в диалог» с действительностью: «задать ей вопрос» и получить ответ, т.е. процесс познания может быть представлен как коммуникативная деятельность. В этом процессе этапы параллельного, в некоторой степени автономного и автопоэтического развития теории и эксперимента, завершаются их симбиозом. Это особое интегративное единство относительно автономных теоретических структур, связанных между собой системой предельных переходов – принципов соответствия [9].

Знания с точки зрения существования – это модель некоторой части реальности в сознании субъекта (так называемые неявные знания) или в научной дисциплине (так называемые явные, обобществленные задокументированные и таким образом объективизированные знания). В отличие от сигналов, которые являются частью физической реальности, знания – это *реальность сознания* субъекта (точнее, *состояние* среды: мозга – биофизической субстанции человека, или вычислительной среды в случае ИИ, которые также являются частями реальности). Знания – это образ, имеющий абстрактно-конкретную природу⁷ и формирующийся путем отражения воспринимаемой части реальности на когнитивное состояние субъекта – его наличное знание, задачи, потребности, мотивы и т.д. Причем знание – это логическая кон-

⁷ «Технологической» основой построения знаний (познания – восприятия/понимания/анализа/синтеза) являются абстракции – категориально-понятийный аппарат, позволяющий определять свойства и вводить аксиомы, формировать модели и теории в конкретных областях.

струкция образов, ассоциирующаяся с действиями, которые выполнялись или могут выполняться в определенной ситуации и которые приводят к определенным последствиям.

С семиотической точки зрения знания – это информация, связанная с контекстом её получения/применения, т.е. **знания** – это *сообщение* (знаки, представляющие содержание) + *используемая понятийная база* (концепт) + *обстоятельства* (коннотат).

С точки зрения формы существования и семантики, знания – это множество фактов⁸, которые:

1) систематизированы, т. е. упорядочены и взаимосвязаны в соответствии с некоторой категориально-понятийной системой, которая отражает основные свойства и взаимодействия, определяющие данную предметную область;

2) опредмечены, т.е. идентифицированы определяемым через характеристические свойства местом в системе (классификации) знаний – предметов, дисциплин;

3) квалифицированы⁹ с учетом области и условий получения/применения, т.е. имеют оценку состоятельности: внутренней полноты и непротиворечивости – теорией, а адекватности и эффективности – практикой или экспериментом;

4) атрибутированы в соответствии со свойствами, определяющими данную предметную область, а также, возможно, именованы.

С точки зрения способа представления, знания – это зафиксированное на носителе, построенное в соответствии с некоторым языком описания/моделирования символическое представление (информационная структура) реальных или абстрактных объектов, явлений, процессов, моделей, теорий, описаний и т.д. Отметим, что в основе языка представления так или иначе лежит категоризация¹⁰ – некоторое множество характеристических признаков (в реальности – множество свойств), а в целом – система признаков. Частный случай – это онтологическая форма представления знаний [11].

С точки зрения формы, знания могут быть представлены (1) в детерминированной символной форме (традиционные тексты, схемы, формулы и т.д.) или (2) в вероятностной форме, отражающей статистические связи данных в совокупности знаний. Аналогично метод построения вывода может иметь схему (1) «по шагам» – путем последовательного «выращивания» результата с контролем промежуточных ре-

зультатов, либо (2) «целостный» – без предъявления промежуточных результатов и оснований для выбора следующих действий и операционных объектов.

Знания могут подразделяться по типу – характеру/области назначения, например, технические знания, знания о продукции (изделии), методологические и т. п., и по виду – способу/носителю знаний: явные/неявные, индивидуальные/коллективные, внутренние/внешние и т.п.).

Знания не являются и не могут быть абсолютными и окончательными, так как это отражение реальности посредством чувств и далее через систему величин.

Знания с точки зрения процесса их построения «рекурсивны»: за основу принимаются факты и утверждения (аксиомы), считающиеся в этот момент достоверными, и далее на их основе возникают построения (теоремы, модели), которые теоретически и/или экспериментально проверяются, порождая новые данные и проблемы, которые, в свою очередь, приводят к пересмотру исходных положений и т.д.

Знания распределены в «пространстве» – между субъектами и во времени – по этапам и формам. Распределенность «во времени» – это возможность передачи знания не только следующим поколениям, но и «самому себе». Последнее важно так как субъект познания развивается, меняется не только состояние его текущего знания и внешние обстоятельства, но и точки зрения, т.е. явное сопоставление с предыдущим знанием является полезным и конструктивным. Это распределенная среда, объединяющая субъектов познания как функционально – организуя деятельность, так и информационно – систематизируя и обеспечивая доступность знаний. Сознание, использующее личностное знание, посредством средств коммуникаций (которые также являются воплощением знаний) взаимодействует с обобществленным знанием, в результате чего формируется новое личное знание (не только практическое и абстрактное, но и эмоциональное – оценочное) и, возможно, пополняется обобществленное.

Если субъектов деятельности больше одного, то такая система должна иметь организацию (а не только структуру), обеспечивающую устойчивые взаимодействия (коммуникации и управление), а если субъекты имеют разные «элементные базы» (например, в случае человеко-машинных систем), то и несколько языков и соответствующих трансляторов/интерфейсов. То есть семантической особенностью деятельности является распределенность и неоднозначность смыслов: объемы и смыслы даже одних и тех же понятий для разных субъектов или разных предметных областей могут не совпадать, поскольку у них может быть разный контекст.

Вследствие распределенности деятельности знания/информация стратифицированы по формам представления (рис. 4). Для неявного знания – это образы-представления ощущения, восприятия, узнавания, именования (кодирования), а для явного (обобществленного) – это категориально-понятийный аппарат, фундаментальные теории, а также самая представительная страта – прикладные знания.

⁸ Факт – это символно зафиксированный акт путем фиксации (отображения, измерения) значения величины или действия, которое воспринимается через свой результат – изменение состояния, представленное величиной. С точки зрения семантической сложности, в [10] факты подразделяются на элементарные (представленные триплетом *сущность-отношение-сущность*), ситуативные, завершенные.

⁹ Именно свойство квалифицированности элементов знаний, точнее его отсутствие в обучающем массиве в явном виде, предопределяет ограниченность статистического (потокового) обучения в искусственном интеллекте.

¹⁰ А в самом начале процесса познания – ощущения/измерения.

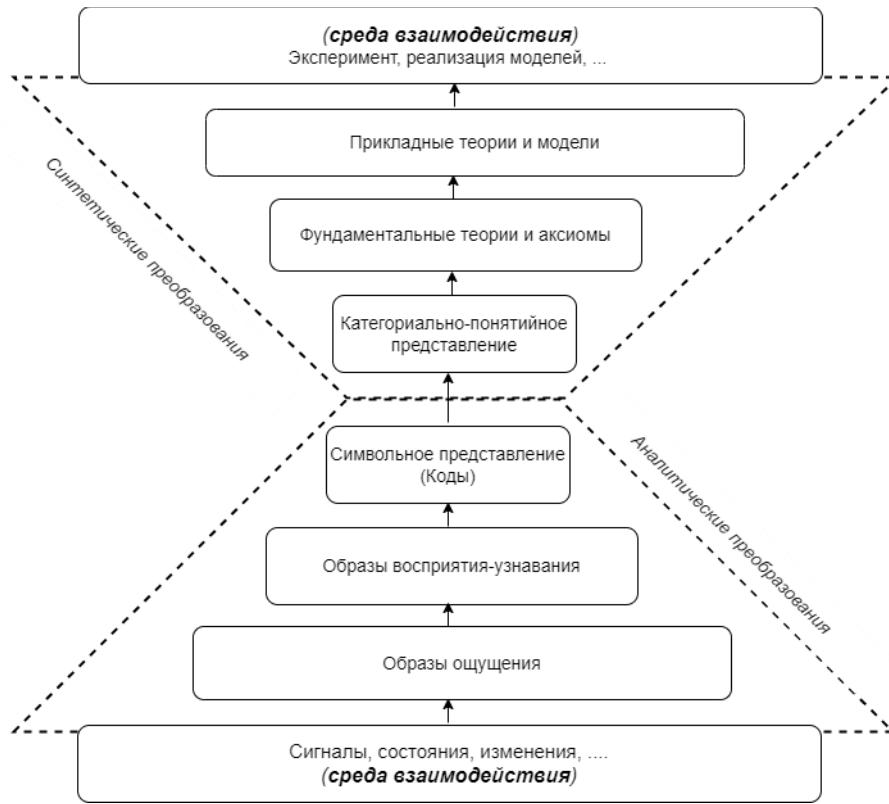


Рис. 4. Страты представления знаний/информации

Стратифицированность имеет следствием:

1) множественность экземпляров отдельного фрагмента (элемента) знания и соответствующего контекста;

2) «категориальное» разделение (сituативно и варианто) компонентов по ролям на объекты и инструменты (реализуя принцип дополнительности), что и обеспечивает возможность анализа (деления по некоторому свойству-признаку) и синтеза знаний (возможно по другому свойству), а также развития деятельности, в том числе познавательной.

Но при этом необходимо отметить, что в искусственном интеллекте (в частности, искусственные нейронные сети) исходными данными для обучения и построения выводов являются преимущественно тексты, относящиеся к многочисленным и разнообразным приложениям – объекты-образы верхней страты, которые количественно практически подавляют (в статистическом смысле) данные других страт. Более того, данные страт индивидуального восприятия-понимания, находящиеся «ближе к реальности» и потому относительно адекватно ее отражающие, составляют явное знание и потому в машинном обучении практически не могут использоваться.

Процесс познания в естественном интеллекте – это не отражение мира (как в ИИ), а построение собственной модели мира. Эта модель включает не только полученные субъектом в процессе деятельности практические и теоретические знания, но и представления этого субъекта о себе и своей роли в мире, о текущей и глобальной цели и т.п. Причем послед-

ние имеют роль «параметра порядка»: их изменения, даже в некоторой части, могут привести к радикальной перестройке сложившейся модели мира. Более того, такая модель – это *некоторая i-я система* (в понимании Общей теории систем Ю.А. Урманцева [12]), т.е. множества объектов, отношений, свойств, выделенные субъектом *условно* – в соответствии с задачей и выбранными им методами.

Знания условны и относительны. И объективны они ровно в той степени, в какой предсказанные принятой моделью свойства и состояния будут соответствовать качеству результатов, которые будут получены в результате взаимодействия с реальностью и которые тоже представлены образами – измерениями в соответствии с некоторой моделью.

Знания фрактально представляют действительность. Они фрагментарны и гранулированы: по своей природе любой результат, любое знание есть аспектный или частный образ, что отвечает принципу дополнительности (в данной предметной области принцип дополнительности/неопределенности означает, что смысл факта/сообщения всегда доопределяется некоторым контекстом: понимание требует использования дополнительных описаний). Причем в случае явной формы существования знаний фрагментарность и гранулированность определяются характером членения действительности и необходимой для прикладной сферы детальностью (точностью и полнотой) представления (модели, описания). Кроме того, «явность» формы требует объяснения и обоснования ее выбора, что само по себе образует новое ассоциированное знание, повышающее полноту и доказательность.

Но именно перечисленные свойства знаний определяют возможность формирования информационного поля для синтеза на основе «старых» знаний новых решений и их оценки.

СООТНОШЕНИЕ ПОНЯТИЙ «ДАННЫЕ», «ИНФОРМАЦИЯ», «ЗНАНИЯ»

Данные, информация, знания – термины широко и по-разному понимаемые, а зачастую воспринимаемые как синонимы. Их содержание и смысловое соотношение определяются, скорее, контекстно – через характер и некоторые особенности области использования.

Например, в области программирования основное понятие – это «данные». Термины же «информация» и «знания» применяются разве что, когда речь идет об использовании результатов работы программы.

В области управления и коммуникаций оперируют преимущественно термином «информация» и, иногда – «данные», причем без конкретизации самих данных, их организации и других «технических» параметров.

В образовании «знания» – основной термин, но практически никак не связанный с понятиями «данные» и «информация».

В науке применяются все три термина, но, в основном, порознь и в разных контекстах.

Отметим, что при такого рода «кузком», специализированном использовании этих терминов практически не возникает необходимости их смыслового согласования и даже их строгого различения. Однако в случае задач анализа/синтеза интегральных систем,

глобального управления, познания в целом (и, в частности, интеллекта) необходимо конструктивно определять эти понятия и их взаимосвязь, т.е. имя понятия должно достаточно четко ассоциироваться со свойствами и методами обозначаемой сущности.

С функционально-ролевой точки зрения информация обычно ассоциируется с актуальным выбором, управляющим воздействием, а данные – с входом, обрабатываемым «сырьем» для «формирования информации». И в этом смысле знания, как результат обработки и накопления информации, уже могут рассматриваться как данность, которая в последующем может стать (точнее, иметь роль) и исходными данными, и информацией, что будет определяться характером их использования.

В [13] динамика соотношений этих понятий целиком представлена с деятельностной позиции: «Любое человеческое знание начинается с накопления фактов, с помощью наблюдения или направленного эксперимента. Но не превращенная в систему, река новых знаний не утолит жажду. Пока хаос новых фактов не структуризован, пока человек не может окунуть взглядом явления в целом, он не может эти знания использовать для практики. Поэтому второй этап – это переработка информации, представление ее в такой форме, которая уже может быть «переварена» человеком. Ну а третий этап – это возвращение к практике, использование знаний для тех целей, ради которых они были созданы».

Применительно к процессу познания (схема которого приведена на рис. 1) эта динамика может быть представлена схемой на рис. 5.

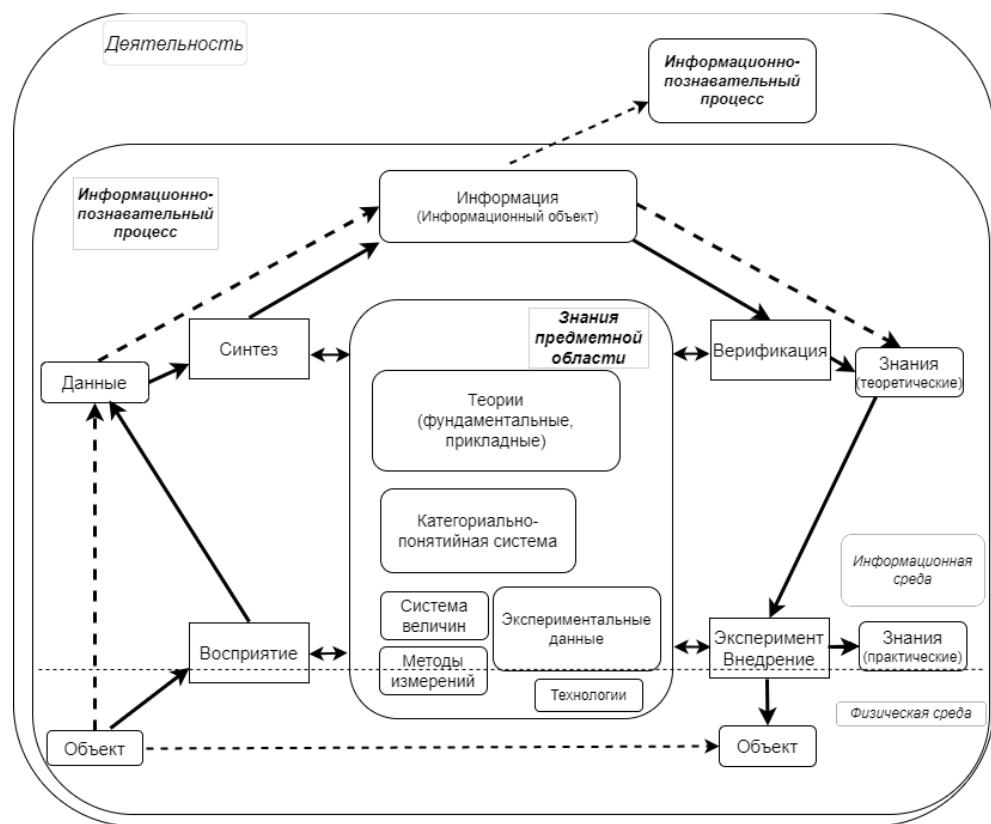


Рис. 5. Динамика и соотношений понятий «информация», «данные», «знания»

Воспринимаемый объект «несет» безусловную (по Д.С. Чернавскому [14]) или связанную (по Л. Бриллюэну [15]) информацию – некоторое «выбранное им» состояние.

Проводя измерение (как и воспринимая органами чувств), т. е. выбирая свойство как величину, присущую состоянию объекта (и, соответственно, единицы измерения), мы получаем (точнее, запоминаем, фиксируем на носителе) уже условную информацию, т.е. «собственная» безусловная информация объекта *обуславливается* выбором способа восприятия/измерения. Это, по терминологии Л. Бриллюэна, есть свободная информация. Но ее свобода ограничена, в частности, конкретностью объекта внимания и набором выбранных свойств (относительно разнообразия свойств, используемых в других задачах и обстоятельствах). Ограничена она и по составу, и по структуре – она односложна, т.е. соответствует понятию «элементарный факт» – символично зафиксированное значение величины, что можно представить тройством <свойство><имеет значение><число единиц измерения>. Именно этот случай структурной разновидности информационных объектов (гранулы) называется «данное» (точнее, «элемент данных»), что соответствует приведенному в ГОСТе [16] определению: данные – это «...объективные факты, которые не могут интерпретироваться вне контекста и дальнейших пояснений». Отметим, что, говоря о семантике данных, надо иметь в виду структуру того, о чем идет речь. Отдельное значение, представленное байтом, строкой и т.п. (т. е. элементарного типа), не имеет семантики. В случае элемента данных, определяемого в программе, мы имеем дело со структурой элементарного факта, где семантика «прилагается» указанием на тип данного, но которая скорее относится к способу представления, чем к смысловому содержанию. Структура данных, поток или массив данных очевидно обладают семантикой. В этом случае семантика задана неявно, но при определенных условиях она может быть выявлена, в предположении, что эти данные являются результатом некоторого конкретного процесса их порождения.

Понятие «Информация» в том же ГОСТе [16] определяется как: «...структурированные данные, обладающие актуальностью и целями, которые могут быть вписаны в контекст, классифицированы, оценены и откорректированы». Можно полагать, что «актуальность» здесь надо понимать не как их сиюминутную применимость и значимость, а как действенность – *потенциальную способность* существенно изменить некоторый фрагмент реальности. А «вписанность в контекст» означает возможное наличие некоторой выделенной («актуальной» в понимании данного ГОСТА) задачи, цели, обстоятельств и т.п., но, очевидно, еще не достигнутой, и потому в процессе ее достижения и задачи, и цели возможно изменятся. При этом информация (точнее, ИнфОб, как макрообразование) представляет собой, обычно, более сложный, композиционный объект, «составленный» из элементарных фактов (в частности, утверждений). Это означает, что уже за счет этого появляется комбинаторность смыслового связывания как внутри, так и вовне информационного объекта (с разными объек-

тами и в разных предметных областях) и, соответственно, вариантность выбора. Информационно-поисковые системы, индексируя тексты, создают условия «динамического хаоса» [14], обеспечивают разнообразие для построения ценных смысловых комбинаций при поиске по запросам: выдается то, что имеет отношение к запросу в каком-либо смысле. Таким образом, информация – это динамическая форма существования знаний [17]. Здесь «динамизм» рассматривается как свойство, предопределяющее возможность получения разных результатов в процессе восприятия одного и того же информационного объекта в зависимости от знаний приемника, т. е., если в случае формирования элемента данных происходит концентрация, то в случае информации происходит рассеяние.

Знания в ГОСТе [16] определяются как «Связанная информация, которая позволяет проводить сравнение, определять степень взаимодействия и принимать решения», а в Международном стандарте [18] – как «Совокупность фактов, событий, убеждений, а также правил, организованных для систематического применения».

Оба определения указывают на конкретизацию информации, смысловое содержание которой до ее применения могло принимать разные значения в зависимости от контекста, т.е. есть «связанность», с одной стороны, обусловлена конкретностью использования и оценки информации, а с другой – системностью знаний: получаемая/синтезируемая информация должна быть «вписанной» в конкретную научную дисциплину как систему знаний. Таким образом в случае знаний происходит концентрация информации путем «составления» только тех фактов, которые имеют отношение к соответствующему смысловому полю (дисциплине, задаче, направлению и т.п.).

И информация, и данные, и знания – это все абстрактно-конкретные образы некоторых объектов-оригиналов. И если рассматривать эти сущности применительно к одному общему оригиналу, то, очевидно, что эти образы будут иметь «общую часть» кодов, отражающую определенные свойства собственно этого объекта. Но их контекстная составляющая будет разной и по содержанию, и по назначению.

Общее соотношение основных понятий представлено схемой на рис. 6.

Субъект деятельности (техническая система или человек) берет данные (сигналы, величины и т.д.) из некоторого, обычно большого, их объема и выбирает из них необходимые не только по величине самих данных, сколько по признаку соответствия контексту, т.е. только те, которые, необходимы для решения конкретной задачи. Далее, для получения прагматического результата данные обрабатываются, т.е. связываются с методом обработки (являющимся одной из форм контекста) и полученный результат (тоже данные) будет связан со своим контекстом – способом его применения, что реально и обеспечит действенность информации для конечного пользователя. Затем практическое использование данных и получаемый результат, в том числе и его оценка, в совокупности формируют знание, которое, будучи опубликованным, становится стимулом или ресурсом (данными) для процессов деятельности.

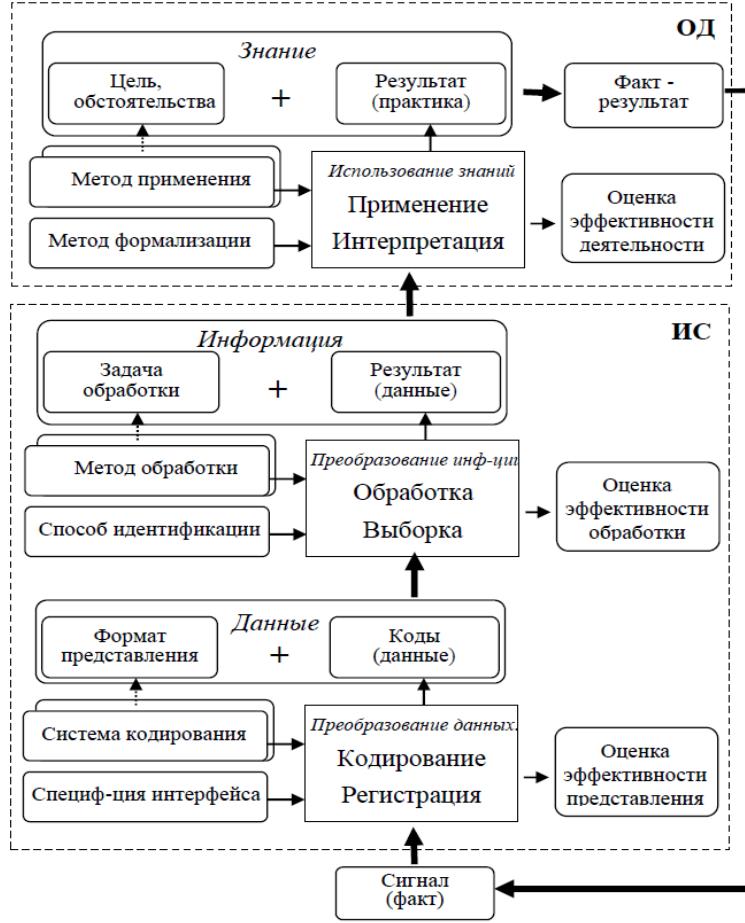


Рис. 6. Контекстное окружение данных.

Здесь ОД – основная деятельность, ИС – информационная система (эта схема приводится по [19]. Она является уточнением и развитием положений [18])

Информационные гранулы *Данные – Информация – Знания* отражают комплексность (контекстную обусловленность) и вложенность объектов, являющихся компонентами обработки данных на соответствующем уровне:

знания = <<<данные & метаданные> & метаинформация> & метазнания>,

где первая пара (операнды в угловых скобках, соединенные оператором «&») – это собственно данные и способ кодирования, вторая – воспринятый текст и контекст, т.е. смысловое содержание, которое интерпретируется или применяется в рамках метазнания, формируя в итоге знания.

При этом по признакам *смысловая конкретность/вариантность и односложность/многосложность информационных объектов*, их соотношение можно иллюстрировать следующим образом:

Смысл\сложность	односложный	многосложный
Вариантный*	Информация	
конкретный	Данные	Знания

*Вариантность смысла обусловлена тем, что результирующее значение информационного объекта (его смысл) будет определяться выбранным контекстом: информация – это выбор из множества возможных состояний, который и определяет значение контекста.

Здесь «информация» может быть представлена как простой конструкцией типа «объект» в рамках объектно-ориентированной парадигмы программирования, так и сложной – типа монографии, учебника и т.п.

Отметим также, что информационные гранулы соответствуют конструктам предметной области следующим образом: данные соответствуют фактам, информация (сообщение) – высказываниям, теоремам и т.д., знания – теориям, моделям, описаниям и т.д.

И подчеркнем, что данные, информация и знания имеют принципиально общую природу, а различаются только в рамках функционально-ролевого подхода.

ИНТЕЛЛЕКТ

Обобщая изложенное, можно сказать, что Интеллект – это логический аппарат, использующий и порождающий знания, обладающий способностью организовывать и контролировать действия в соответствии с некоторой моделью, а также формировать и оценивать эту модель. Интеллект – следствие: (1) взаимодействия (в том числе информационного типа и с обратной связью) с окружающей средой; (2) многоуровневого представления знаний в локальной (субъекта) и внешней памяти; (3) использования языка как средства представления, коммуникаций и моделирования.

Систему искусственного интеллекта по аналогии со схемой архитектуры ЭВМ фон Неймана [20] –

«**вычислительная машина=процессор+память**» и с аналогичной обобщенной структурой программы Н. Вирта [21] – «**программа=алгоритм+данные**» можно рассматривать как «**система ИИ=алгоритмы_вывода+знания**». Причем, как отмечалось в основополагающих работах по программированию, такое деление на типы достаточно условно: программы (точнее инструкции) существуют в виде данных и, соответственно, могут быть модифицированы в процессе их выполнения (что является принципиальной основой саморазвития), а методы построения вывода (решения) – это знания, зафиксированные в виде данных, которые также могут изменяться в процессе их использования. Процессор и память – это аппаратная среда, элементы которой различаются по функциям, но имеют общее назначение – сохранять или изменять состояние элементов, т.е. практически и по существу конструкты **ЭВМ – Программа – Система ИИ** представляют собой в совокупности стек вложенных объектов (точнее, «матрешку»). Соответственно, система искусственного интеллекта – это программная реализация логических методов на данных в вычислительной среде. Причем, как естественный интеллект в своей основе – это физико-химические процессы и структуры, так и искусственный интеллект – это информационные процессы, использующие структуры данных. Отметим также, что этот стек подобен стеку **Данные – Информация – Знания** [6, 18].

В этом смысле искусственный интеллект представляет собой третичные модели реальности, т.е. это модели, построенные на вторичных моделях (знаниях, представлениях, описаниях), основанных, в свою очередь, на первичных моделях – ощущениях, проявляющихся в нервной системе при воздействии стимула, который единственно и является реальностью. А знания – это образ мира (части реальности, включая познающего), сформированный на мире образов в соответствии с некоторым образом. Причем, стратифицированность и опосредованность представлений (использование образов, а не оригиналов реальности) предопределяет необходимость исследования и доказательства адекватности выводимого знания. Это, в свою очередь, требует соответствующих инструментов и знаний (паттернов, правил, критериев и т.д.). Причем этот процесс также обусловлен и целями, и назначением, и другими факторами, т.е. тем, что является внешним по отношению к интеллекту.

В контексте разграничения слабого и сильного искусственного интеллекта, можно сказать, что слабый ИИ (точнее, **узкий, специализированный ИИ**) соответствует понятию «рассудок». А в том случае, если система искусственного интеллекта будет обладать способностями (1) избирательно взаимодействовать с окружающей средой и другими агентами ИИ (что вполне реализуемо в цифровой среде), (2) выявлять проблемные ситуации, а также формировать цели, проявлять «любопытство» и «интерес» к поиску более выгодных условий существования, то это будет **сильный ИИ**. А если он еще будет обладать и

формировать¹¹ систему внутренних ценностей, а также будет способен к самовоспроизведению и самосовершенствованию на основе хранимых, синтезируемых и/или находимых вовне знаний, то это будет **общий искусственный интеллект (AGI)**¹². Таким образом можно сказать, что AGI – это система ИИ, способная к рациональной (т.е. в соответствии с определенными искусственным интеллектом целями, доступными ему ресурсами и принятыми критериями) деятельности (т.е. целенаправленно взаимодействующий с окружающей средой и ресурсами знаний), в том числе и к выявлению проблемных ситуаций (и, возможно, их ожиданию, т.е. моделированию возможных), а также к динамическому (в процессе деятельности) изменению выбранных целей и парадигм. Эта спецификация близка к понятию **разум**¹³, которое, обобщая многочисленные определения, можно определить как высший тип мыслительной деятельности человека, его способность мыслить всеобще (интеллект), способность к анализу, абстрагированию и обобщению (рассудок), а также способность целенаправленно извлекать/воспринимать, синтезировать, передавать и *реализовывать* знания, а также их квалифицировать, систематизировать и реорганизовывать в процессах деятельности и приспособления к окружающей среде.

Отметим, что все компоненты и функции естественного интеллекта, представленные на приведенных схемах (см. рис. 1-3), так или иначе, в том или ином составе уже практически осуществляются. В том числе возможна реализация (точнее, пока симуляция) и эмоций, и потребностей, и целей, хотя их адекватность глобальным целям организма-социума сомнительна. По существу, это уже **искусственная жизнь**, которая должна будет взаимодействовать и конкурировать с другими формами жизни, в том числе с человеком и обществом. Но отсюда следует, что смысл жизни и, соответственно, ее нормативы надо было бы определить до реального зарождения этой жизни.

О СООТНОШЕНИИ ИСКУССТВЕННОГО И ЕСТЕСТВЕННОГО ИНТЕЛЛЕКТА

Естественный интеллект, рассматриваемый как составляющая процесса деятельности, как реализующий ее саморазвивающийся механизм, может быть представлен в виде иерархической системы распределенных взаимодействующих систем (периферийная и центральной нервной системы, понятийные системы, физиологические системы, системы коммуникаций, социальные системы и т.д.). Эти системы рассредоточены (1) по средам функционирования: организм (физиче-

¹¹ Причем, в условиях понимания относительности знаний и бренности существования их носителей.

¹² Или, согласно [22], – **автономный ИИ**, обладающий адаптивностью, эмоциональным аппаратом и свободой принятия решений.

¹³ Различие разума и рассудка как двух “способностей души” есть уже в античной философии: если **рассудок** как низшая форма мышления познает относительное, земное и конечное, то разум направляет на достижение абсолютно-го, божественного и бесконечного.

ская среда – тело, нервная система), ментальная среда (психика, сознание), коммуникации (язык, каналы), окружающая среда и социум (реализации/результаты, эффекты), и (2) по стратам представления: ощущения, узнавание, понимание, анализ-синтез, реализация/оценка. При этом для каждого состояния (страта-среда) есть (1) свое операционное пространство – множество объектов в некоторой области памяти; (2) свои критерии (пороги) успешности результата, причем чем выше сложность страты, тем больше факторов будет в составе критерия; (3) свои правила выполнения итераций, необходимых для доводки решения, причем критерий остановки будет учитывать не только необходимость/значимость результата, но и мотивацию, и волю субъекта в достижении цели.

Мультимодальность представлений (как следствие разнородности подсистем и сред) предопределяет, что один воспринимаемый объект порождает некоторое разнообразие образов, создавая возможность явно или неявно проводимого многоаспектного анализа и использования нескольких факторов. Их разноплановость обеспечивает большую полноту при поиске, выборе и оценке, а также подготовку альтернатив.

Иерархичность совокупной системы деятельности предполагает, что подсистемы имеют разные цели и, соответственно, используют разные фрагменты знаний, которые, возможно, по-разному представлены и по-разному упорядочены. Следовательно, необходимо либо провести координацию целей – изменить их (и, возможно, методы решения), причем, возможно не только локальные, но и глобальные – всего организма (индивида и/или общества), либо выбрать одну из них, подавляя цели всех остальных подсистем.

В случае естественного интеллекта такая координация/выбор – это неявные процессы (а может быть одноактный процесс, если рассматривать его с точки зрения квантово-волнового подхода), которые на логическом уровне практически не контролируются, которые трудно анализировать и верифицировать. Но на практике они приводят обычно к адекватному результату, что происходит, вероятно, вследствие мультимодальности и гранулированности представления знаний, запаса неявных (в том числе отрицательных) знаний, а также диверсификации потребностей, мотивов и целей, но, главное – к их гармоничному *естественно* сложившемуся сочетанию.

В случае же искусственного интеллекта цели, задачи и используемые знания сформированы изначально, однозначно и в хорошо определенной форме, причем исходя из целесообразности и ожидаемой полезности (для человека и задачи), а также предполагаемого пути получения результата.

ВЫВОДЫ

Целесообразно сформировать рабочую «принципиальную» (логическую) схему искусственного интеллекта и определить возможные архитектуры (как совокупность существенных, определяющих свойств, а не состав и структуру, как это часто встречается в

современных публикациях) в совокупности с возможными областями их применения и классами задач.

Для устойчивого развития искусственного интеллекта необходимо формально определить, что есть знания (и, желательно, стандартизовать в совокупности с определением ИИ), их типы, виды, формы представления и, соответственно, языки определения и манипулирования, а также методы использования, в том числе для обучения систем ИИ.

Надо определить стратегии развития и применения искусственного интеллекта и, соответственно, формы и технологии представления знаний, а в итоге установить идти ли по пути создания одного глобального СуперМозга (и, может быть единого хранилища информации), или *Системы взаимодействующих разнородных специализированных систем ИИ*. Но можно заметить, что уже вследствие третичности ИИ (достоверность выводов, за исключением тривиальных или формальных так называемых «прописных истин», всегда будет под сомнением и будет требовать дополнительных действий, связанных с выходом за пределы системы) напрашивается распределенная форма.

И, очевидно, надо сохранить самостоятельные естественные формы хранения/поиска знаний (библиотеки, информационные системы и т.д.), которые имеют в своем составе не только массив документов/записей, но и средства систематизации знаний (классификации, указатели, тезаурусы и т.д.). Действительно, характеристические свойства знаний – систематизированность, предметность и квалифицированность предопределяют, что любой элемент состоявшегося или нового знания должен соотноситься с некоторой другой устойчивой сущностью: научными дисциплинами, фактами, классификациями и т.д. Здесь «устойчивая сущность» означает отдельность ее существования в виде самостоятельного объекта, а «устойчивость» надо понимать не в физическом аспекте (как неизменяемость), а скорее в статистическом, т.е. семантика (существо, смысл) элемента знаний может и должна существовать во многих экземплярах (как информационные объекты): в разных видах и формах, с разной полнотой и может быть на разных языках. Кроме того, важно понимать, что «существовать» – значит взаимодействовать: каждый экземпляр элемента знаний порождает собственную смысловую цепочку, приводящую к новому знанию. При этом формируется семантическое поле, создавая, в частности, основу и для анализа свойств этого элемента знаний, и для его рекурсивной актуализации.

Кроме того, одним из условий и источников ускоренного развития Знаний выступает явная, обобщенная форма их представления, обеспечивающая параллельное и разноплановое использование отдельных элементов (утверждений, данных, теорий и т.д.). При этом разные методы, критерии и технологии фактически порождают совокупность оценок (своебразную мультимодальность), сравнительный анализ которых позволяет объективизировать знания, а они по определению всегда субъективны. Это, в

свою очередь, позволяет судить о степени достоверности этих оценок. Значимы и аналогии – схожие схемы построения вывода, которые могут быть характерны как для научно-технической, так и для художественной литературы. Следует отметить, что и ошибочные теории/опыты имеют определенную конструктивность, в частности, не только отсекая возможно ошибочные пути развития, но и формируя ассоциации в семантическом поле проблемы.

Сохранение знаний и информации – это не только и не столько очевидная функция обеспечения *неразрушения их носителей* (точнее, *неизменяемость состояния*), сколько поддержание системной составляющей (ее актуализация и гармонизация – изменение, отражающее развитие науки и практики), а также средств коммуникации – языков и их носителей. Именно языки в условиях постоянного изменения представлений потребителей об объектах и предметах действительности обеспечивают необходимую доступность информации, которая еще не утратила полезность, но не будет найдена вследствие произошедших со временем изменений в языке предметной области.

СПИСОК ЛИТЕРАТУРЫ

1. Черниговская Т.В. Чеширская улыбка кота Шрёдингера: мозг, язык и сознание. – Москва: Издательство АСТ, 2022. – 496 с.
2. ГОСТ Р 59276—2020 Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения. – Москва: Стандартинформ, 2020.
3. Максимов Н.В. Когнитивность информационного поиска в контексте информационности знания // Научно-техническая информация. Сер.1. – 2022. – №11. – С.1-17; Maksimov N.V. Cognitive Nature of Information Search in the Context of Informativeness of Cognition // Scientific and Technical Information Processing. – 2022. – Vol. 49, № 4. – P. 238-254.
4. Бейтсон Г. Разум и природа: Неизбежное единство. – Москва: УРСС: КД «ЛИБРОКОМ», 2016. – 256 с.
5. Бурцев М.С., Бухвалов О.Л., Ведяхин А.А. Сильный искусственный интеллект: На подступах к сверхразуму. – Москва: Интеллектуальная литература, 2021. – 232с.
6. Симонов П.В. Мотивированный мозг. – Москва: Наука, 1987. – 287 с.
7. Дорожкин А.М. Научный поиск как постановка и решение проблем. – Н.Новгород: Нижегородский гуманитарный центр, 1995. – 108 с.
8. Богданов А.А. Тектология: Всеобщая организация науки. – Москва: Финансы, -2003. – 496 с.
9. Аршинов В.И., Буданов В.Г. Когнитивные основания синергетики // Синергетическая парадигма. Нелинейное мышление в науке и искусстве. – 2002. – С. 67-108. – URL: <https://spkurdyumov.ru/networks/kognitivnye-osnovaniya-sinergetiki-arshinov> budanov/?ysclid=ll5i03t8vl763832099_(дата обращения 08.06.2020).
10. Максимов Н.В., Голицына О.Л. От семантического к когнитивному информационному поиску. Основные положения и модели глубинного семантического поиска // Научно-техническая информация. Сер. 2. – 2022. – № 6. – С. 1-15; Maksimov N.V., Golitsyna O.L. From Semantic to Cognitive Information Search: The Fundamental Principles and Models of Deep Semantic Search // Automatic Documentation and Mathematical Linguistics. – 2022. – Vol. 56, № 3. – P. 145-159.
11. Максимов Н.В., Лебедев А.А. Онтологическая система «знания-деятельность» // Онтология проектирования. – 2021. – Т. 11, № 2(40). – С. 185-211.
12. Урманцев Ю.А. Общая теория систем: Состояние, приложения и перспективы развития // В сб. «Система, Симметрия, Гармония». – Москва: Мысль, 1988. – С. 38-124.
13. Моисеев Н.И. Предисловие к сборнику «Число и мысль». – Москва: Знание, 1977. – 176 с.
14. Чернавский Д.С. Синергетика и информация (динамическая теория информации). – Москва: Едиториал УРСС, 2004. – 288 с.
15. Бриллюэн Л. Научная неопределенность и информация. – Москва: Книжный дом «ЛИБРОКОМ», 2019. – 272 с.
16. ГОСТ Р 57321.1-2016 Менеджмент знаний. Менеджмент знаний в области инженеринга. Часть 1. Общие положения, принципы и понятия. – Москва: Стандартинформ, 2017.
17. Максимов Н.В. Информация и знания: природа, концептуальная модель // Научно-техническая информация. Сер. 2. – 2010. – № 7. – С. 1-10; Maksimov N.V. Information and Knowledge: Nature and the Conceptual Model // Automatic Documentation and Mathematical Linguistics. – 2010. – Vol. 44, № 4. – P. 177-186.
18. ISO/IEC 2382-1 Information technology – Vocabulary – Part 1: Fundamental terms. – ISO, IEC, 1993.
19. Голицына О.Л., Максимов Н.В., Попов И.И. Информационные системы и технологии. – Москва: ФОРУМ: ИНФРА-М, 2014. – 400 с.
20. von Neumann J. First Draft of a Report on the EDVAC. – University of Pennsylvania, 1945.
21. Вирт Н. Алгоритмы + структуры данных = программы. – Москва: Мир, 1977. – 406 с.
22. Жданов А.А. Автономный искусственный интеллект.– 4-е изд. – URL: <https://postnauka.org/books/38231> (дата обращения 10.05.2023).

Материал поступил в редакцию 13.10.23.

Сведения об авторах

МАКСИМОВ Николай Вениаминович – доктор технических наук, профессор, профессор Национального исследовательского ядерного университета МИФИ (НИЯУ МИФИ), Москва
e-mail: NV-MAKS@YANDEX.RU

О.В. Сюнтуренко, Н.В. Червинская

Информационное обеспечение науки и промышленности Китая*

Показаны амбициозные задачи информационной политики Китая. Рассматривается структура научных исследований и разработок, а также вопросы управления и финансирования. Отмечается, что основные вложения в научно-исследовательскую и инновационную деятельность в Китае осуществляют предприятия. Показаны ключевые информационные ресурсы науки и промышленности Китая, а также базовые элементы информационной инфраструктуры научно-исследовательской и промышленной сферы. Приводятся фактические и сопоставительные данные. Рассмотрены некоторые значимые факторы негативного характера, препятствующие ускоренному развитию экономики Китая.

Ключевые слова: информационная политика, инновации, фундаментальные исследования, информационные ресурсы, кадровая политика, патентная активность, информационная инфраструктура, сетевые технологии, суперкомьютеринг, информационная безопасность

DOI: 10.36535/0548-0019-2024-01-2

ЦЕЛИ И СТРУКТУРА ЗАДАЧ ИНФОРМАЦИОННОЙ ПОЛИТИКИ КИТАЯ

Развитие национальной системы информационного обеспечения науки, образования и промышленности Китайской Народной Республики ориентировано на создание государства инновационного типа – мощной опоры для превращения Китая в одного из мировых лидеров в сфере науки и технологий. Идеи повышения качественного уровня китайской национальной информационной инфраструктуры как основы развития научно-технического и промышленного потенциала страны получили свое дальнейшее институциональное воплощение (и конкретную реализацию) в 14-м пятилетнем экономическом плане развития на 2021–2025 г. [1]. Основная цель этого плана состоит в том, чтобы как можно быстрее сократить зависимость от иностранных технологий и импортных ресурсов, а также ускорить существующие планы промышленной модернизации и технологических инноваций. Приоритетные задачи: 1) решение проблем энергетической (низкоуглеродное развитие Китайской экономики), водной и экологической безопасности; 2) перевооружение промышленности посредством передовых информационных технологий и технологий в области новых материалов; 3) развитие сельского хозяйства и промышленности с помощью биотехнологий; 4) ускоренное развитие морских и

космических технологий, 5) развитие фундаментальных и передовых прикладных исследований.

Основная задача Китая – в обозримой перспективе стать мировым лидером в области науки и технологий. Китай обладает одним из самых больших научно-исследовательских комплексов в мире, который стремительно вырос за последнее десятилетие. По многим количественным и качественным параметрам НИОКР, страна находится в пятёрке мировых лидеров. Одной из главных причин быстрого развития науки являются масштабные финансовые вливания. Они стали возможны благодаря как значительному увеличению бюджетных трат, так и активности промышленного сектора.

В настоящее время в КНР наблюдается процесс перехода от ресурсного типа экономического развития к инновационному. Формируется новая инновационная модель развития экономики, являющаяся следствием перехода к доминированию таких элементов как информационные и коммуникационные технологии, внешний и внутренний трансферт инноваций, развитие рынка инноваций и интеллектуальной собственности. Инновационный процесс – это единственный в своем роде процесс, объединяющий науку, технику, экономику, предпринимательство и управление. Правительство страны делает ставку на ускоренную коммерциализацию научных достижений и ведущую роль рынка в определении актуальности и важности научных исследований. Важным направлением реформ является повышение качества управления наукой и эффективности НИОКР.

* Материал настоящей статьи заслушивался и обсуждался на Международной конференции LIBCOM в г. Суздале в ноябре 2023 г.

Таким образом, государственная инновационная политика Китая является одной из важнейших составляющих всего комплекса проводимой в стране государственной информационной политики в научно-промышленной сфере.

СТРУКТУРА НАУЧНЫХ ИССЛЕДОВАНИЙ И РАЗРАБОТОК. УПРАВЛЕНИЕ И ФИНАНСИРОВАНИЕ

В настоящее время научно-исследовательской и инновационной деятельностью в Китае занимаются 3650 научно-исследовательских институтов (НИИ), включая учреждения Китайской академии наук (около 440 институтов, лабораторий, проектных центров и станций), Китайской академии общественных наук (76 институтов и научных центров); 2560 университетов с 11,7 тыс. научно-исследовательских подразделений. Действует также система провинциальных академий, подведомственных местным правительству. Помимо этого, в Китае инновационными разработками различного рода занимаются более 73 тыс. промышленных предприятий [2]. И именно предприятия, в рамках государственной научно-технической политики, являются ключевой силой развития науки. В структуре научных исследований и разработок экспериментальные работы значительно превалируют над теоретическими исследованиями: порядка 70% валовых внутренних расходов приходится на научные разработки, 6% – на фундаментальные исследования, а остальная часть расходов – на прикладные исследования. Фундаментальную науку полностью поддерживают государство. В 2017 г. на долю фундаментальных исследований приходилось около 25% всех расходов центрального правительства на науку.

Китайский феномен объясняется и определяется в основном тремя факторами: высокой квалификацией специалистов различного уровня на передовых промышленных предприятиях; сильным инженерным корпусом и, самое главное, – высокой мотивацией максимизации прибыли со стороны руководства и коллектива предприятия. Еще в начале восемидесятых годов XX в. руководители отраслей зачастую выделяли предприятию исходного сырья меньше запланированного на 8% (например), а электроэнергии на 6%, а если предприятие выполняло план на 107%, то эти 7% продукции оставались в полном распоряжении коллектива и дирекции. Такая технико-экономическая политика в масштабах страны позволила в относительно сжатые сроки осуществить эффективную трансформацию экономического мышления и производственных процессов без обычного в таких случаях системного провала эффективности.

Основные вложения в научно-исследовательскую и инновационную деятельность в Китае осуществляют предприятия. Эта особенность сформировалась еще до середины 90-х гг. ХХ в. (когда появились соответствующие статистические данные) и усилилась в последнее десятилетие. В 2016 г. на НИОКР предприятий приходилось 77,5% всех затрат на науку и инновации в стране. Расходы на науку повышает неизменно и государство. Уже в 2015 г. в номинальных значениях траты всех уровней госорганов на

НИР составили 87,1 млрд долларов США (по паритету покупательской способности), заняв второе место в мире по вложениям в науку и составив 72% от расходов мирового лидера в государственных расходах – США. Доля расходов на науку в бюджете Китая в 2016 г. достигла 4,1% [2] (в России этот показатель составляет ~ 0,9–1,1%).

Рост ВВП Китая по годам с 2017 г. по настоящее время выглядит следующим образом:

- 2017г. – 12 трлн 062,2 млрд долл.
- 2018г. – 13 трлн 407,3 млрд долл.
- 2019г. – 14 трлн 216,5 млрд долл.
- 2020г. – 14,86 трлн долл.
- 2021г. – 17,76 трлн долл.
- 2022г. – 18,10 трлн долл.
- 2023г. – 19,37 трлн долл. (прогнозная оценка)

Средние темпы экономического роста – 8,1%.

В качестве сопоставления следует отметить, что ВВП России в 2019 г. составлял примерно 1 трлн 623 млрд долл. США, что примерно соответствует годовому ВВП Испании (1 трлн 474 млрд).

За период с 2016 по 2019 гг. объем расходов Китая на НИОКР вырос на 11,8%, что значительно превышает показатели таких технологически развитых стран, как США (7,3%) и Япония (0,7%) [3].

Значительные успехи отмечаются во всей системе науки Китая: в университетах, научно-исследовательских институтах и на предприятиях. Китайская академия наук неизменно лидирует в рейтинге *Science* [4–6], который отражает публикационную активность организаций в 68 ведущих журналах по ключевым областям науки, а также по количеству статей, индексируемых на платформе *Web of Science*. Последние семь лет Китай опережает США по количеству статей, представленных в базе данных *Scopus* [5]. Высокие показатели достигаются в том числе и активным международным сотрудничеством – каждая пятая статья китайских учёных опубликована в международном соавторстве. С 2010 г. Китай – лидер по количеству заявок на регистрацию патентов, большая часть которых исходит от предприятий.

Ключевой исполнительный орган, а также ведомство, которое вырабатывает основные положения научно-технической политики и конкретные законодательные акты, – это Министерство науки и техники Китая (МНТ). С 2008 г. Госсовет возложил на МНТ обязанность выступать координирующей структурой в процессе реализации научно-технической политики. В марте 2018 г. в ходе реформы Госсовета это министерство было усилено через слияние его с Государственным управлением по делам иностранных специалистов и Комитетом государственного фонда естественных наук (ГФЕН) – ключевым грантодателем в КНР. До слияния МНТ имело наибольшие финансовые ресурсы в науке, а вместе с деньгами ГФЕН оно стало главным распорядителем бюджетных ассигнований на все виды исследований и разработок.

Огромное внимание в Китае уделяется кадровой политике в области науки и техники. Только по официальным данным за последние двадцать лет Китай послал на учебу за рубеж свыше 300 тыс. студентов. В их число не входят аспиранты и учёные, принявшие гранты различных зарубежных фондов, а также

студенты, обучающиеся за границей на собственные деньги. Для прошедших обучение за рубежом и вернувшихся на родину студентов, государство создает специальные исследовательские центры, укомплектованные новейшим экспериментальным оборудованием и приборами. Следует подчеркнуть, что благодаря всем рассмотренным аспектам научно-технической политики, процесс развития высокотехнологичного сектора экономики Китая происходит эволюционно и имеет комплексный характер. Такой подход сыграл важную роль в развитии передовых отраслей, в увеличении количества наукоемких предприятий, в диверсификации экономики, в росте ВВП и в целом – в повышении конкурентоспособности экономики Китая.

КЛЮЧЕВЫЕ ИНФОРМАЦИОННЫЕ РЕСУРСЫ НАУКИ И ПРОМЫШЛЕННОСТИ КИТАЯ

Информация и услуги в рамках глобального национального информационного проекта CNKI (China National Knowledge Infrastructure). Важнейшая задача этого проекта – тотальная оцифровка ресурсов знаний Китая, а также создание платформы для глобального распространения услуг с добавленной стоимостью (CNKI позиционирует себя как главную базу знаний КНР) [7]. В 2019 г. был запущен CNKI 2.0, который более полно использует Интернет, большие данные и искусственный интеллект (CNKI – <https://ru.abcdef.wiki/wiki/CNKI?ysclid=liulgvd8148970931>). В стране пользователями CNKI являются: 100% вузов и научных организаций; 95% учреждений начального образования; 60% правительственный организаций; 60% общественных организаций; 80% армейских и других силовых структур; 40% промышленных предприятий.

В интегрированную базу данных входят следующие ресурсы:

полнотекстовая база данных периодики Китая. К 3563 важнейшим периодическим изданиям доступ открывается прежде, чем очередной номер выходит на бумажном носителе;

десять электронных тематических коллекций по 168 отраслям знания. Особое внимание уделяется математике, физике, химии, экологии, сельскому хозяйству, медицине и здравоохранению, информатике, электронике, экономике и менеджменту;

полнотекстовая база защищённых в Китае докторских и магистерских диссертаций. Ресурс содержит 310 тыс. докторских диссертаций, защищенных (начиная с 1984 г.) в 411 различных организациях, а также три с лишним миллиона магистерских диссертаций, защищённых в 635 образовательных учреждениях;

полнотекстовая база данных докладов на важнейших конференциях в КНР. Включает в себя 2,13 млн материалов 17 333 конференций, начиная с 1999 г.;

полнотекстовая база данных докладов на международных конференциях. Ресурс содержит более 700 тыс. материалов 5826 международных конференций (в том числе за пределами КНР) начиная с 2010 г.;

БД ежегодных изданий Китая. Ресурс предоставляет доступ к государственным годовым статистическим справочникам по провинциям, городам, уездам и поддерживает поиск актуальной статистически ин-

формации, которая охватывает 18 основных отраслей народного хозяйства КНР;

базы данных на английском языке. Включают: периодические издания (134 наименования, 240 тыс. статей); доклады и другие материалы конференций (2366 томов, 273 тыс. статей), периодические издания на двух языках (128 томов, 3251 статья);

специальные профессиональные издания. Этот раздел CNKI состоит из двух «площадок знаний»:

первая – содержит три базы данных: о китайских целебных травах, о диагностике и методиках лечения, а также о рецептах китайской медицины. Кроме этого, медицинская площадка содержит цифровую библиотеку по четырём специализированным темам: онкология, остеология, оториноларингология и гипертоническая болезнь,

вторая – посвящена законодательству КНР и объединяет три базы данных: лучшие докторские и магистерские диссертации в области юридических наук, судебная и иная правовая практика, закон и право. Ресурс, помимо законодательных и нормативно-правовых актов КНР, аккумулирует более 200 периодических изданий в области юриспруденции и законодательства, более шести тысяч статей из периодических профильных изданий.

По данным Similarweb в мае 2018 г. сайт CNKI посетили 25,5 млн пользователей. Англоязычная версия сайта насчитывает 500 тыс. ежемесячных посещений.

Ресурсы китайского Центра развития информации. Этот Центр создан в 2000 г. в Пекине при Министерстве информатики КНР для обеспечения стратегического планирования и исследовательской деятельности в производстве электронных средств информации. Он является авторитетной организацией Китая, ведущей исследования в области изучения рынка информационных технологий, тестирования ИТ-продуктов, печатных СМИ и интернет-СМИ, предоставляющей информацию и консультации в этой сфере.

Ресурсы китайского Центра оборонной научно-технической информации. Роль этого Центра заключается в сборе, переводе с иностранных языков, обобщении, анализе, публикации и распространении материалов, относящихся к вопросам национальной безопасности страны. В банк данных Центра стекается научно-техническая информация со всего мира – отчёты по конференциям, научные публикации, рекламная литература и любая другая техническая информация. В нем создана огромная библиотека, в которой есть военные журналы практически всех стран мира, имеется также Бюро переводов. Центр обеспечивает информацией все заинтересованные организации оборонного научно-технического комплекса Китая.

Использование ресурсов платформы Web of Science. Компания Clarivate Analytics в партнерстве с Китайской академией наук разместила Китайскую базу данных научного цитирования на платформе Web of Science. Более 1200 журналов и 5,2 млн записей с 1989 г. – это огромный источник библиографической информации, цитат и ссылок на статьи в крупнейших научно-инженерных журналах, опубли-

кованных в КНР. Китайская база данных научного цитирования стала первым продуктом не на английском языке, доступном на платформе *Web of Science*. Теперь в *Web of Science* реализована поддержка стандарта кодирования Юникод, благодаря чему расширились возможности поиска (искать информацию можно не только на английском языке), а пользователям Китайской базы данных научного цитирования и других ресурсов *Web of Science Core Collection* стали доступны новые варианты пользовательской настройки.

Несколько лет назад, по сообщения журнала «*Nature*», платформа *Web of Science* была продана своим владельцем, компанией Thomson Reuters, в рамках реализации активов общей стоимостью \$ 3,55 млрд долл. Согласно условиям сделки, пакет акций поделен между Onex Corporation, базирующейся в Торонто (Канада), и Baring Private Equity Asia со штаб-квартирой в Гонконге (КНР) [8].

В 2016 г. по данным, опубликованным в журнале «*Nature*», Китай впервые опередил США по количеству статей, отраженных в базе *Scopus* [5]. В последние годы на учёных вузов или НИИ КНР приходилось 15-20% всех статей, проиндексированных в базах *Web of Science* и *Scopus*.

В 2022 г. Китай обогнал и впервые вышел на первое место в мире в элитном рейтинге Топ 1% наиболее высокоцитируемые научные публикации в сфере естествознания. На долю китайских авторов пришлось 27,2%, или 4744, из 1% самых цитируемых работ в мире. Это больше, чем у США (24,9%, 4330 публикаций). В рейтинге учитывались публикации в восьми научных областях: химия, материаловедение, физика, информатика/математика, инженерия, окружающая среда / геонауки, клиническая медицина и фундаментальные науки о жизни.

Национальные ресурсы патентной информации. С 2010 г. Китай – лидер по количеству заявок на регистрацию патентов, большая часть которых исходит от предприятий. Патенты в области искусственного интеллекта (ИИ) являются основой развития системы искусственного интеллекта в стране. За период с января 2018 г. – по октябрь 2021 г. было подано в общей сложности 650 тыс. патентных заявок на ИИ в более чем 100 странах и регионах по всему миру, из которых три страны с наибольшим количеством заявок – это Китай, США и Япония. Количество заявок в Китае составляет 445 тыс., или 68,5%, количество заявок в США – 73 тыс., или 11,2%, количество заявок в Японии – 39 тыс., или 6,0%. Сегодня Китай стал крупнейшей страной по количеству заявок на патенты в области искусственного интеллекта, намного обогнав США. Стоит отметить, что за последние четыре года темпы роста заявок на патенты и темпы внедрения инноваций в области искусственного интеллекта в Китае всегда были в 1-2 раза выше, чем в США, а темпы высоки. Китай занимает первое место в мире по количеству патентных заявок на искусственный интеллект, а по общему объему патентных заявок – первое место в мире семь лет подряд [9].

БАЗОВЫЕ ЭЛЕМЕНТЫ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ И ПРОМЫШЛЕННОЙ СФЕРЫ

В этом разделе мы рассмотрим три наиболее значимых сегмента современной китайской информационной инфраструктуры.

Национальная система телекоммуникаций, сетевые технологии и глобальная сеть Интернет. Развитие этого сегмента влечет смену парадигмы функционирования системы информационного обеспечения исследований и разработок – от иерархической к сетевой. Создание распределенных сетевых информационных ресурсов (ИР) – наиболее бурно развивающееся направление информатизации научно-промышленной сферы. Ожидается, что быстрое развитие высокоскоростных телекоммуникационных технологий 5G будет способствовать ускорению процесса новой индустриализации Китая. В 2021 г. количество зарегистрированных промышленных интернет-предприятий в Китае достигло 620 тыс. [10]. Официальные данные свидетельствуют, что по состоянию на конец марта 2023 г. в Китае было построено более 2,6 млн базовых станций 5G, а количество владельцев смартфонов с поддержкой этого стандарта связи превысило 620 млн. 5G-сети покрывают всю территорию страны. Согласно отчету Ассоциации GSM (GSMA), технология 5G в КНР обгонит 4G в 2024 г. и станет самым распространённым мобильным стандартом в стране [10,11].

Национальная система информационной безопасности. Одна из важнейших задач научно-технической политики Китая – это обеспечение цифрового суверенитета. Нормативный механизм правового регулирования критической информационной инфраструктуры и обеспечения ее безопасности характеризуется взаимосвязанностью и отражает общий характер режима цифровой политики КНР. Закон о кибербезопасности страны устанавливает общие нормы, подзаконные акты – специальные нормы, а стандарты содержат высокотехнологические методические рекомендации, которые могут прояснить возможную неоднозначность общих и специальных норм.

Технически за государственный контроль Интернета в Китае отвечает система «Золотой щит», первая версия которого была введена в эксплуатацию по всей стране в 2003 г. Итоговая версия доработана к 2011 г. Сегодня «Щит» включает несколько подсистем: а) управления безопасностью; б) информирования о правонарушениях; в) контроля ввода и вывода; г) мониторинга; д) управления трафиком.

Суперкомпьютинг. Высокопроизводительные вычисления превратились в КНР в важную отрасль, обслуживающую все основные направления науки и экономики. Практически все крупные компании используют суперкомпьютеры как для повышения своего научно-производственного потенциала, так и для повышения экономической эффективности принимаемых решений. Суперкомпьютеры используются при планировании, в режиме оперативного управления производственными процессами, при управлении большими техническими системами, при моделировании систем

и процессов. Суперкомпьютерные технологии начинают применять также при прогнозировании погоды, мониторинге землетрясений, планировании транспортного движения и решении задач защиты окружающей среды.

В мировой список Топ-500 за ноябрь 2020 г. было включено в общей сложности 212 суперкомпьютеров китайского производства [12] (что составляет почти 40 % всех компьютеров этого списка).

Следует отметить, что в последнее время темпы развития суперкомпьютеров в Китае замедлились из-за технологического давления США.

Несколько отклоняясь от основной темы статьи необходимо кратко остановится на некоторых острых проблемах развития китайской экономики национального масштаба.

A. Острый (и нарастающий) дефицит пресной питьевой воды практически на всей территории Китая. Водоснабжение промышленных предприятий и объектов социальной сферы в значительной степени осуществляется за счет артезианских скважин. Однако в подавляющем большинстве случаев качество воды низкое (в основном из-за повышенной засоленности). В приморских районах Китая (Шанхай) довольно широко используются оросительные установки морской воды. Однако все они базируются на высокозатратных энергетических технологиях и в целом отличаются недостаточно высоким качеством пресной воды.

B. По оценкам европейских специалистов для комфорtnого проживания населения пригодно примерно не более 6% территории Китая (горы, пустыни, полупустыни), поэтому усиление комплексной работы по предотвращению опустынивания земель имеет огромное значение (особенно северных, северо-западных, и северо-восточных регионов). Например, в столице Китая Пекине в зимние месяцы при минус 11 градусов мороза и полном отсутствии снега песок из пустыни Гоби может хрустеть на зубах и засыпать улицы. С 2012 г. Китаю удалось сократить площадь пустынных земель почти вдвое (24,5 млн га). Создана 41 комплексная образцовая зона по предотвращению опустынивания, а также 99 национальных пустынных парков. На севере страны построили экологический барьер против ветра и песка с научно обоснованным выбором посадочного материала, расположением деревьев, кустарников и трав. За последние 10 лет в рамках проекта «Три Севера» площадь лесопосадки выросла до 33,3 млн га. Было обработано 336 тыс. км² опустыненных земель, а среднее количество дней с песчаными бурями с 6,8 снизилось до 2,4 в год [13, 14].

C. Демографические проблемы Китая. Реализация социально-экономических преобразований происходит в сложных условиях развития демографической ситуации в Китае. По официальным данным населения страны составляет более 1,4 млрд. человек. В результате проведения государственной демографической политики появился сильный дисбаланс между рождаемостью и смертностью. Число пенсионеров намного превышает численность молодежи. В ближайшем будущем убытки государства по материальному содержанию, выплатам пенсий, выдаче бес-

платных медицинских препаратов превысят налоги от зарплат работающих лиц. Исследования показывают, что при сохранении нынешних тенденций рождаемости к 2050 г. основу населения страны составят лица пенсионного возраста (от 60 до 90 лет) [15].

ЗАКЛЮЧЕНИЕ

В последнее десятилетие во многом благодаря эффективной деятельности Национальной системы информационного обеспечения Китай совершил большой скачок вперед в сфере научных исследований и образования. Серьёзные успехи отмечаются во всей системе науки – в университетах, НИИ и на предприятиях.

Результативность комплексного информационного обеспечения в стране иллюстрирует пример из области природопользования, экологии и борьбы с опустыниванием (Гоби).

Необходимо подчеркнуть, что развитие высокотехнологичного сектора экономики Китая происходит эволюционно и имеет комплексный характер. Такой подход, включая развитие современной системы информационного обеспечения, сыграл важную роль в развитии передовых отраслей, в увеличении количества научноемких предприятий, в диверсификации экономики, в росте ВВП и благополучия народа и в целом в повышении конкурентоспособности экономики Китая.

В заключение следует отметить, что в условиях мирового экономического спада, для сохранения нынешних темпов развития национальной экономики Китая, необходимо решение актуальной и безотлагательной проблемной задачи – формирования емкого платежеспособного внутреннего рынка вследствие относительно низкого уровня жизни большинства населения Китая. Для правительства КНР обеспечение перехода от экспортно ориентированной экономики последних десятилетий к новой экономической модели безусловно сложная комплексная задача. Однако можно констатировать, что научно-технологические и экономические предпосылки для успешного решения этой задачи у современного Китая есть.

СПИСОК ЛИТЕРАТУРЫ

1. 14-я пятилетка Китая. Новостной портал о Китае PRC.today – Китай сегодня. – URL: <https://prc.today/14-ya-pyatiletka-kitaya/?ysclid=lk013597w2368838124> (дата обращения 03.08.2023).
2. Иванов С. А. Научно-техническая политика Китая: приоритеты догоняющего развития и результаты (cyberleninka.ru). 2018. – URL: <https://cyberleninka.ru/article/n/nauchno-tehnicheskaya-politika-kitaya-prioritetnye-dogonyayuscheho-razvitiya-i-rezulatty/viewer> (дата обращения 23.08.2023).
3. СИНЬХУА Новости. Расходы Китая на НИОКР в 2020 году превысили 2 439 млрд юаней_ Russian.news.cn. – URL: http://russian.news.cn/2021-09/23/c_1310203264.htm (дата обращения 24.08.2023).
4. Institution outputs // Springer Nature. – URL: <https://www.natureindex.com/institution-outputs/generate/>

- All/global/All/weighted_score (дата обращения: 15.05.2018).
5. Fei Shu. Comment to: Does China need to rethink its metrics and citation-based research rewards policies? // Scientometrics. – 2017. – Vol. 113, № 2. – P.1229-1231.
 6. Tollefson J. China declared world's largest producer of scientific articles // Nature. – 2018. – Vol. 553, № 7689. – P. 390.
 7. Национальная инфраструктура знаний Китая | Digital Russia (d-russia.ru). – URL: <https://d-russia.ru/natsionalnaya-infrastruktura-znanij-kitaya.html?ysclid=lilz83m32p46989374> (дата обращения 24.08.2023).
 8. Web of Science будет продана / Научная Россия (электронный журнал) 13.07.2016 (scientific-russia.ru). – URL: <https://scientificrussia.ru/articles/web-of-science-budet-prodana> (дата обращения 24.08.2023).
 9. Объем патентных заявок Китая занимает первое место в мире семь лет подряд. – URL: https://translated.turopages.org/proxy_u/en-ru.ru.c42c7c12-64c27b4d-040e1ebc-74722d776562/https://pandaily.com/chinas-patent-application-volume-ranks-first-world-seven-consecutive-years/ (дата обращения 27.07.2023).
 10. Статистический бюллетень индустрии связи Китая за 2021 год. – URL: https://www.miit.gov.cn/jgsj/yxj/xxfb/art/2022/art_3b457a2cda504fe89b75605fe7235492.html (дата обращения 25.08.2023).
 11. Связь (рынок Китая). – URL: Связь (рынок Китая) (tadviser.ru) (дата обращения 25.08.2023).
 12. Китай добился серьёзных успехов в создании суперкомпьютеров • Новостной портал о Китае PRC.today – Китай сегодня. – URL: <https://prc.today/kitaj-dobilsya-seryoznyh-uspehov-v-sozdaniisuperkompyuterov/?ysclid=1knoimrlv4458000520> (дата обращения 29.07.2023).
 13. Китай превращает пустыни в цветущие сады // Комсомольская правда 28 июня – 05 июля 2023. – С. 33.
 14. Как Китай превращает свои пустыни в леса. План правительства работает. | Мир и деньги | Дзен (dzen.ru). – URL: https://dzen.ru/a/X_WP7rsU1U_7m-ld (дата обращения 25.08.23).
 15. Демографическая проблема Китая (spravoch-nick.ru). – URL: https://spravochnick.ru/gosudarstvennoe_imunicipalnoe_upravlenie/demograficheskaya_politika/demograficheskaya_problema_kitaya/?ysclid=lpjyk_szdds326283278 (дата обращения 30.11.23).

Материал поступил в редакцию 01.12.23.

Сведения об авторах

СЮНТИОРЕНКО Олег Васильевич – доктор технических наук, профессор, главный научный сотрудник ВИНИТИ РАН
e-mail: olegasu@mail.ru

ЧЕРВИНСКАЯ Надежда Викторовна – ВРИО директора Всероссийского института научной и технической информации РАН
e-mail: chervin@viniti.ru

ОРГАНИЗАЦИЯ ИНФОРМАЦИОННОЙ РАБОТЫ

УДК [001.891:002]:316.77

О.Н. Шорин

Коммуникационные процессы в научной среде

Рассматриваются основные направления развития коммуникационных процессов, обеспечивающие доступ к результатам научных исследований. Анализируются причины недоступности научных публикаций для ученых, приводится описание процесса становления идеи открытого доступа, дается обзор текущего положения дел в опубликовании учеными своих достижений в открытом доступе. Представлены результаты изучения статистики, наглядно демонстрирующие основные тренды в области научных коммуникаций в мировом масштабе.

Ключевые слова: коммуникационные процессы в науке, открытый доступ, научные журналы, публикации, централизованная подписка, открытые репозитории, библиометрия, авторское право, цитирование, Web of Science

DOI: 10.36535/0548-0019-2024-01-3

ВВЕДЕНИЕ

Обмен знаниями между учеными, открытое обсуждение применяемых методов, теорий и моделей, независимая верификация полученных результатов – это одни из основополагающих принципов развития науки. Изначально эти функции выполняли профессиональные сообщества ученых, в которых исследователи могли открыто дискутировать с коллегами на различные темы, связанные с той или иной теорией, экспериментом, сформулированными выводами. С развитием книгопечатания такие дискуссии перекочевали в научные издания – сообщества распространяли журналы, в которых печатались статьи о различных достижениях науки, и любой заинтересованный исследователь мог с ними ознакомиться, попытаться подтвердить изложенные в них положения или опровергнуть их, используя такой же механизм публикации научных статей.

Подобная практика получения, подтверждения или опровержения, накопления научного знания сохранилась до сих пор. Более того, со временем у управленицев, администрирующих финансирование различных областей науки, появилось желание иметь возможность каким-то образом измерять развитие научных направлений количественными показателями [1]. В результате появилась концепция, утверждающая, что конечным продуктом исследования ученых являются опубликованные научные статьи. Таким образом, потребность ученых в ознакомлении с научными статьями по интересующим их направлениям, а также возможность печататься в научных

журналах стали необходимыми условиями для развития научного знания.

Соглашаясь с постулатом о том, что наука не имеет границ и знания накапливаются в интересах всего человечества, становится очевидным, что для устойчивого развития науки необходимо обеспечить доступ к научным публикациям абсолютно всех ученых вне зависимости от их национальной принадлежности, политических взглядов, финансовых возможностей и пр.

На протяжении большого периода времени возможность доступа ученых к научным публикациям обеспечивали библиотеки, приобретая периодические издания и предоставляя их своим читателям. Однако во второй половине XX в. начался необратимый процесс коммерциализации рынка научной периодики, когда такие коммерческие издательства, как *Springer* (Германия) и *Elsevier* (Нидерланды), стали выкупать права на научные журналы у небольших организаций. Выкупленные журналы они объединяли в «пакеты», предлагая библиотекам приобретать эти пакеты целиком, что приводило к тому, что библиотеки были вынуждены выкупать не только те издания, в которых они заинтересованы, но и остальные, входящие в «пакет». Это неминуемо вело к росту стоимости подписки. Такой подход по увеличению размеров «пакетов» и их стоимости получил название *Big Deal* (Большая сделка) [2]. В частности, согласно исследованиям [3], цена подписки на научные журналы в США с 1984 по 2001 гг. выросла на 651,6% для журналов по зоологии, на 614,0% для журналов по химии и физике, на 578,6% для журналов по медицине.

Возникла парадоксальная ситуация, которую эксперты называют «двойной оплатой», когда государство выделяет деньги на исследования, результатом которых являются публикации в научных журналах, а затем государство вынуждено закупать эти журналы у коммерческих компаний для обеспечения учебных актуальной информацией [4].

С широким распространением высокоскоростного подключения к сети Интернет ведущие издательства вместо «пакетов» с печатными версиями журналов предлагают оформление подписки на доступ к электронным версиям журналов. Библиотеки и научно-исследовательские учреждения стали объединяться в консорциумы для приобретения подписок на нужные им «пакеты» в интересах консорциума. Это объяснялось следующей логикой: чем больше размер консорциума, тем ниже стоимость подписки для отдельно взятого члена этого консорциума. Масштабируя такую логику, в конечном итоге, это может привести к появлению единого оператора в стране, который будет закупать право доступа к подписке в интересах всех научных организаций в стране.

С 2005 г. в рамках федеральных целевых программ по поддержке приоритетных направлений развития научного и технологического комплекса Российской Федерации Минобрнауки РФ осуществляет целевое финансирование подписки на зарубежные электронные ресурсы. Исполнителями государственных контрактов по предоставлению доступа к зарубежной научной информации в разные периоды были назначены [5]:

- в 2005–2013 гг. – консорциум НЭИКОН;
- в 2014–2019 гг. – Государственная публичная научно-техническая библиотека (ГПНТБ) России;
- с 2020 г. по настоящее время – Российский фонд фундаментальных исследований (РФФИ), который постановлением Правительства Российской Федерации от 29 июля 2022 года № 1357 был переименован в Российский центр научной информации (РЦНИ) ¹¹.

Согласно отчетам РЦНИ стоимость лицензионных и сублицензионных соглашений по доступу к электронным изданиям и научным информационным ресурсам в интересах российских ученых в 2020 г. составила 3 555,88 млн руб. [6], в 2021 г. – 3 733,7 млн руб. [7], в 2022 г. – 4 918,87 млн руб. [8].

Эти данные показывают, что затраты государства на обеспечение доступа к электронным ресурсам увеличиваются год от года. К тому же, приобретая доступ к изданиям, а не сами издания, появляются новые риски, которые в юридической литературе называются «концом владения» [9], – доступ к тем или иным подписанным электронным ресурсам закрывается в зависимости от различных факторов: например, от геополитической ситуации, от наложенных одними странами на другие санкциями, от отсутствия финансирования и т.д.

Такое положение вступает в противоречие с изначальными принципами развития науки в рамках всего

человечества: открытость, безграничность, всеобъемлющий и беспрепятственный доступ к актуальной научной информации. К тому же большинство научных исследований финансируется государством, поэтому результаты этих исследований, в частности научные публикации в журналах, являются общественным достоянием и должны быть доступны всем желающим. В сложившейся ситуации возникло и получило развитие движение за обеспечение открытого доступа (Open Access) к научным публикациям [10].

Журнальный кризис, безусловно, оказал решающее воздействие на развитие идей открытого доступа, но, помимо этого, специалисты выделяют еще несколько типов причин: экономические, управленческие, развитие электронных публикаций и новых сетевых технологий, а также социальные мотивы [11].

ОТКРЫТЫЙ ДОСТУП

Принято считать, что точкой отсчета в движении за открытый доступ к научным публикациям является Будапештская инициатива «Открытый доступ», которая была принята 14 февраля 2002 г. по итогам конференции, созданной фондом «Институт Открытое Общество» [12]. Инициатива предлагала использовать современные технологии, в частности сеть Интернет, для безвозмездной публикации учеными плодов своих трудов для беспрепятственного распространения новых знаний.

Будапештская инициатива была направлена, в первую очередь, на предоставление доступа всем желающим к журнальным статьям, прошедшим экспертную оценку. Но она предлагала публиковать в открытом доступе и другие материалы, например, нерецензированные препринты, с помощью которых ученые могли бы быстрее обнародовать свои результаты и получать замечания от коллег.

Для достижения этой цели предлагалось использовать два взаимодополняющих инструмента [13]:

1) создание открытых архивов, в которых ученые могли бы свободно размещать свои публикации, документы, данные и т.п. Подразумевалось, что вместе с этим будет разработан программный инструментарий, позволяющий обмениваться информацией между архивами, агрегировать метаданные, осуществлять поиск по распределенной системе открытых архивов;

2) создание нового поколения альтернативных журналов, нацеленных на распространение знаний без каких-либо ограничений. Для финансового обеспечения таких журналов предлагалось использовать пожертвования частных лиц и фондов, привлекать финансирование от университетов и государства, развивать систему дополнительных платных сервисов, оказываемых журналами на основе научных публикаций.

На данный момент Будапештская инициатива открытого доступа подписана 1 573 организациями и 6 765 физическими лицами.

22 октября 2003 г. в развитие идей Будапештской инициативы открытого доступа [14] была опубликована Берлинская Декларация об открытом доступе к научному и гуманитарному знанию, в которой сформулированы требования, предъявляемые к публикациям открытого доступа, а также действия, направ-

¹ Постановление Правительства Российской Федерации от 29.07.2022 № 1357 "О федеральном государственном бюджетном учреждении "Российский центр научной информации". – URL: <http://publication.pravo.gov.ru/Document/View/0001202208020006?rangeSize=1> (дата обращения: 28.07.2023).

ленные на поддержку дальнейшего развития принципов открытого доступа.

Берлинская Декларация первоначально была подписана 19-ю представителями различных научных и исследовательских организаций [15]. Они договорились о том, что Берлинская конференция, посвященная вопросам открытого доступа, будет проходить ежегодно. На Берлинской конференции, которая состоялась 6-7 июня 2023 г., всеми присутствовавшими на заседании делегациями, было принято заявление, состоящее из следующих пунктов [16]:

1. Глобальный переход к открытому доступу должен продвигаться гораздо более быстрыми темпами.

2. Неравенство несовместимо с научными публикациями.

3. Академическое самоуправление является императивом в научных публикациях.

4. Выбор автора по предоставлению набора прав, а также авторские права на публикацию должны быть полностью поддержаны юридическими инструментами и не ограничиваться издательствами.

Одним из аргументов ученых в пользу обеспечения открытого доступа к их научным публикациям является утверждение, что статьи, опубликованные в открытом доступе, цитируются чаще по сравнению со статьями, опубликованными в подписных журналах. Часть исследований подтверждает это предположение [17]. Однако есть ученые, которые приходят к противоположным выводам [18]. Дело в том, что в различных областях науки устаревание информации происходит с разной скоростью, из-за чего подходы к использованию цитат варьируются в зависимости от тематической специализации [19]. Поэтому на результат исследования очень сильно влияет выбор временного интервала, на протяжении которого анализируется цитируемость публикаций. Некоторые ученые предполагают, что у статей, опубликованных в открытом доступе, на первоначальном этапе цитируемость выше, но со временем этот показатель выравнивается [20], поскольку принято считать, что статьи, опубликованные в журналах, доступных по подписке, более авторитетны, и из-за этого на них ссылаются чаще [21].

Данные, доступные с помощью открытого доступа

В Будапештской инициативе открытого доступа изначально декларировалось, что необходимо развивать два направления: архивы открытого доступа и научные журналы нового типа для публикации статей [22]. Впоследствии выяснилось, что для управленцев, администрирующих распределение финансовых ресурсов по различным областям научных исследований, огромный интерес представляет доступ к библиометрическим базам данных, в которых хранится информация о цитировании научных статей одними учеными другими [23]. В связи с этим, имея в виду данные, доступные с помощью открытого доступа, можно выделить три основных их типа:

1) научные архивы, в которых, в основном, размещаются нерецензированные публикации и сопутствующие данные;

- 2) библиометрические базы данных;
- 3) публикации в научных журналах, которые прошли рецензирование.

Научные архивы

Публикация статьи в научном журнале – это долгий и трудоемкий процесс, поскольку в высококачественных журналах статья проходит строгое рецензирование (из рассмотрения исключены так называемые «хищнические» журналы, которые готовы за определенную плату оперативно опубликовать статью без рецензирования или с использованием саморецензирования [24]). Чаще всего рецензирование занимает несколько месяцев, иногда этот срок может увеличиваться до нескольких лет [25]. К тому же на этапе рецензирования иногда могут отсеиваться до 70% поступающих в редакцию рукописей [26]. В связи с этим ученым необходим механизм, с помощью которого они могут оперативно публиковать полученные результаты. Таким механизмом выступают публично доступные открытые архивы научной информации – репозитории открытого доступа.

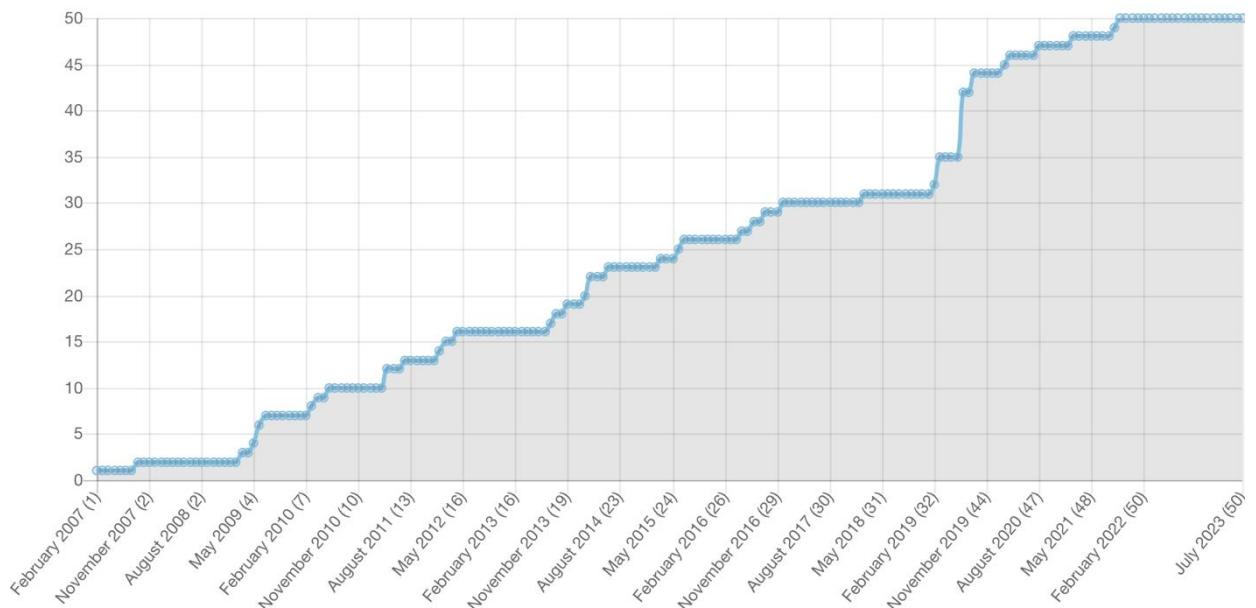
В таких репозиториях ученые могут размещать препринты, чтобы закрепить за собой ведущую роль в определенном направлении исследований, еще не прошедшие рецензирования и потому не опубликованные статьи (депонирование), данные, необходимые для воспроизведения и верификации своих исследований, аудио- и видео-материалы, которые невозможно опубликовать в научных журналах.

Традиционно, репозитории принято делить на [27]:

- институциональные – репозитории, созданные научно-исследовательскими организациями, для размещения в них материалов своих сотрудников. Такие репозитории создаются учреждениями для подтверждения значимости института, а также для учета и контроля проводящихся в организации научных исследований;
- тематические – репозитории, накапливающие материалы по определенной тематике вне зависимости от труда у устройства ученого;
- общего назначения – такие репозитории не ограничены какой-либо тематикой и не связаны с институтами.

Инициативой «Открытые архивы» (*Open Archives Initiative – OAI*) разработаны стандарты, которым должен соответствовать репозиторий для обеспечения совместимости взаимодействия с другими системами, что позволяет обеспечить эффективное распространение информации, а также создание сервисов, основанных на контенте, хранящемся в научных архивах. Один из основных стандартов этой инициативы – протокол для сбора метаданных – (*Protocol for Metadata Harvesting – OAI-PMH*) [28]. По этому протоколу можно агрегировать информацию о данных, хранящихся в различных репозиториях, и организовывать поиск с последующим доступом к данным вне зависимости от того, в каком репозитории они физически расположены, тем самым реализуется гибридная модель функционирования научных архивов: распределенное хранение с выделенными центральными узлами, осуществляющими индексирование и поиск по различным репозиториям.

Growth of OpenDOAR



Динамика роста репозиториев России в *OpenDOAR* в 2007 – 2023 гг.

В настоящее время разработано несколько программных платформ, с использованием которых можно организовать и администрировать открытый репозиторий. Большинство таких программных платформ соответствует стандартам *OAI* (*Open Archives Initiative*), поэтому тот, кто хочет организовать научный архив, может выбрать ту платформу, которая больше всего ему подходит. Среди наиболее распространенных программных платформ можно выделить: *DSpace*, *Eprints*, *WEKO*, *Digital Commons* [29].

Существуют также два известных реестра открытых репозиториев, предназначенных для регистрации научных архивов, сбора информации о них и распространения этой информации среди научного сообщества: *Registry of Open Access Repositories (ROAR)* [30] и *Directory of Open Access Repositories (OpenDOAR)* [31]. На 21 июля 2023 г. в ROAR содержится информация о 4 725 репозиториях, а в OpenDOAR – о 6 036. По данным ROAR в России 67 репозиториев, по данным OpenDOAR – 50. Динамика роста репозиториев России в OpenDOAR приведена на рисунке.

По данным OpenDOAR, две трети (66%) репозиториев в России работает под управлением программной платформы *DSpace*; 41 репозиторий содержит информацию по социальным наукам, 38 – по медицинским, 36 – по гуманитарным наукам, по математике и технологическим исследованиям, 35 – по искусству и инженерному делу [32].

В 2011 г. было проведено исследование об отношении ученых к открытым репозиториям научной информации [33], которое показало, что ученые предпочитают узнавать об исследованиях своих коллег из открытых научных архивов, но при этом не

сильно стремятся к тому, чтобы выкладывать свои наработки в репозитории. Отчасти это связано с тем, что ученые более высоко ценят свои публикации в научных журналах, которые прошли рецензирование, по сравнению с той информацией, которую они самостоятельно делают доступной без предварительного рецензирования. Вследствие этого, репозитории остаются на вторых ролях по отношению к публикациям в научных журналах, но имеют потенциал для развития в качестве площадки обмена научными данными, которые не учитываются в отчетности, но представляют интерес для информационного насыщения среды [34].

Библиометрические базы данных

В последнее десятилетие усилиями Министерства науки и высшего образования Российской Федерации и различных фондов, выдающих гранты на исследования, в нашей стране была построена система оценки научного труда, которая в значительной степени опиралась на библиометрические показатели отдельно взятых сотрудников и организаций в целом. В соответствии с этой системой принято учитывать различные показатели: количество публикаций, количество соавторов, количество аффилий, индекс Хирша, импакт-фактор (квартиль, дециль, процентиль и т.п.) журнала, в котором размещена публикация, количество ссылок на опубликованные статьи, отсутствие самоподтверждений и т.п. Все эти показатели отбирались из двух зарубежных баз данных – *Web of Science* (владелец компания *Clarivate*) и *Scopus* (владелец издательство *Elsevier*), которые были доступны всем научным организациям России через механизм национальной и централизованной подписки.

По поводу этих коммерческих баз данных, которые принадлежат зарубежным компаниям, периодически высказывалась критика из-за их сосредоточенности, в основном, на журнальных статьях и недостаточном охвате других научных публикаций, но методики оценки деятельности организаций, докториальных советов, заявок на гранты, отчетов, разработанные Минобрнауки России, Высшей аттестационной комиссией, грантовыми фондами основаны на использовании именно этих баз данных, поскольку в 2012 г. был подписан указ Президента РФ от 7 мая 2012 г. № 599 "О мерах по реализации государственной политики в области образования и науки", в соответствии с которым необходимо было обеспечить «увеличение к 2015 году доли публикаций российских исследователей в общем количестве публикаций в мировых научных журналах, индексируемых в базе данных "Сеть науки" (WEB of Science), до 2,44 процента»².

Использование базы данных *Web of Science* в качестве основного источника библиометрической информации было практически безальтернативным. Однако, в мае 2022 г. доступ к этой базе ученым из России был закрыт, а в январе 2023 г. прекратился доступ и к базе данных *Scopus* [35]. Таким образом вопрос об использовании открытых библиометрических систем встал более остро.

Необходимо упомянуть, что в 2010 г. появился проект *OpenCitations*, основной целью которого было сделать возможным свободную публикацию пристайных списков литературы в формате, удобном для машинной обработки [36]. Основным достижением этого проекта является создание, поддержка и наполнение открытого репозитория данных о цитированих в научных публикациях. На первоначальном этапе подготовки этого репозитория были использованы научные статьи по медицине, которые располагались в открытом доступе. Впоследствии было разработано программное обеспечение, позволяющее автоматизировать процесс наполнения этого репозитория. С апреля 2017 г. *OpenCitations* стала одним из шести основателей инициативы «Открытое цитирование» (*Initiative for Open Citations – I4OC*), основная цель которой заключается в обеспечении доступности данных о цитированих с учетом принципов структурированности, открытости и отделимости [37].

Структурированность означает, что данные о каждой публикации и её цитированих представлены в машиночитаемом формате, таким образом эта информация становится доступной для обработки программными средствами. Принцип отделимости означает, что данные о цитировании доступны и могут быть проанализированы без необходимости доступа к текстам самих статей, в которых эти цитированиа упоминаются. Открытость подразумевает, что данные о цитировании доступны всем безо всяких ограничений и могут быть использованы для создания сервисов более высокого уровня.

² Указ Президента РФ от 7 мая 2012 г. N 599 "О мерах по реализации государственной политики в области образования и науки" // Российская газета. – 09.05.2012. – № 102.

Отчасти благодаря такому подходу к использованию данных и наработкам одних систем для создания новых сервисов в последние двадцать лет появилось некоторое количество полиграфических библиографических баз данных, отличающихся друг от друга и охватом, и типами хранимых данных, и предоставляемым функционалом. В частности, можно выделить следующие системы: *Google Scholar* (год основания – 2004), *AMiner* (2006), *Lens* (2013), *Scilit* (2014), *Semantic Scholar* (2015), *Microsoft Academic* (2016), *Dimensions* (2018), *scite_* (2019), *MyRA* (2019), *Exaly* (2022), *OpenAlex* (2022). Согласно обзору функциональных возможностей и наполнению этих баз данных, такие системы как *Lens*, *Dimensions* и *OpenAlex* содержат информацию о списках цитирований, что, как уже отмечалось, важно для проведения библиометрического анализа [38].

Исследования показывают, что уже в 2021 г. объем данных о цитированих, доступных с использованием систем на основе открытых данных, достиг паритета по сравнению с данными из систем *Web of Science* и *Scopus* [39]. Таким образом в мире появилась бесплатная открытая альтернатива проприетарным системам, с помощью которых можно решать различные библиометрические задачи, в частности проводить оценку публикационной активности отдельно взятых сотрудников, коллективов авторов или целых организаций и учреждений.

Публикации в рецензируемых журналах

Рост цен на подписку издательства объясняют тем, что, помимо затрат на типографические услуги для распространения печатной продукции и на поддержание сетевой инфраструктуры для подписки на электронные версии журналов, они вынуждены тратить огромные средства на рецензирование статей, поступающих в редакции. И поскольку количество статей растет из года в год [40], то и затраты издательств для обеспечения качества публикуемых статей увеличиваются, что компенсируется за счет роста стоимости подписки.

Труд рецензентов, а также коллективов редакционных коллегий журналов должен быть оплачен. И если издательство опубликует все журналы в открытом доступе, лишившись тем самым возможности получить деньги за продажу печатной версии или за обеспечение доступа к электронной версии, то оно будет вынуждено зарабатывать деньги каким-то иным способом.

Один из таких способов – оказание услуг по переводу статей на другие языки для публикации в так называемых «переводных журналах». Существует предположение, что статью на английском языке прочитает, а, соответственно, и процитирует гораздо большее количество человек, чем ту же самую статью, например, на русском языке. В связи с этим параллельно с журналами на русском языке издательства иногда создают точно такие же журналы на английском языке, и берут с авторов плату за перевод статьи с русского на английский язык и публикацию переведенной статьи в англоязычной версии журнала. А поскольку перевод статьи оплачивает автор, который либо получает зарплату от государства, либо

использует для этого деньги полученного им гранта, то вместо уже упомянутой проблемы двойной оплаты за публикацию статьи и доступа к ней возникает проблема тройной оплаты: государство оплачивает труд ученого, который публикует статью, перевод этой статьи на иностранный язык для переводной версии журнала и подписку на доступ к электронной версии журнала, в котором в итоге будет опубликована переведенная статья.

Отработав механизм получения денег от авторов за перевод их статей на другие языки, в условиях открытого доступа издательства вполне логично переносят бремя расходов с потребителей своей продукции (читателей) на поставщиков контента для своих журналов (авторов) [41]. Существуют, конечно, издательства, которые не берут денег ни с читателей, ни с авторов (так называемый «бриллиантовый» или «платиновый» тип открытого доступа [42]), но часто издательства предлагают авторам оплатить услуги за обработку публикации (*Article Processing Charge – APC*) в обмен на то, что их статья будет опубликована в открытом доступе [43]. Стоимость такой услуги в высокорейтинговых журналах достаточно высока – несколько тысяч долларов за одну статью [44]. Получается, что эту сумму государство вынуждено заплатить за то, чтобы статья попала в открытый доступ.

А что такое открытый доступ? Будапештская инициатива трактует открытый доступ как такой доступ, при котором разрешается свободное чтение статьи, а также её вторичное использование: индексирование содержимого, выгрузка статьи программными средствами, использование контента для любых других законных целей. Существуют и другие определения, которые трактуют открытый доступ лишь как возможность читать научную литературу онлайн [45]. Третий тип определения открытого доступа требует, чтобы статьи, размещенные в открытом доступе, были цифровыми, доступными онлайн и бесплатными [46].

Пользуясь тем, что диапазон формулировок понятия открытый доступ достаточно широк, издательства требуют от авторов подписать собственные варианты лицензионных соглашений, в которых четко прописывается кто и какими правами обладает на публикуемую статью. Специалисты условно делят эти лицензионные соглашения на несколько типов [47]:

- Libre – пользователь может читать статью, а также вторично её использовать: индексировать, скачивать программными средствами, архивировать, а также работать с текстом любым законным способом;
- Gratis – пользователь может только читать статью онлайн;
- золотой – статья печатается в так называемом «открытом журнале» – журнале, в котором пользователь может читать все статьи на сайте издательства;
- зеленый – статья публикуется в журнале, который доступен за плату, но при этом автору разрешается самостоятельно разместить эту статью в открытом репозитории;
- бронзовый – статья доступна на сайте издательства, но при этом отсутствует какой-либо документ, явно описывающий режим ее использования;

• гибридный – статья публикуется в журнале, который доступен за плату, но при этом издательство размещает ее в свободный доступ под открытой лицензией;

- отложенный – этот тип похож на «гибридный» тип, но с одним отличием – издательство размещает статью в открытый доступ не моментально, а по прошествии определенного времени, которое называется периодом эмбарго.

Исследования показывают [47], что в том или ином типе открытого доступа располагается 27,9% статей. При этом большинство из этих статей опубликовано с использованием «бронзового» типа – 58%, «гибридный» тип используется у 12,9% статей, опубликованных в открытом доступе, «золотой» тип используется у 11,5% статей, а «зеленый» тип – у 17,2%.

Для того чтобы упорядочить рынок публикаций открытого доступа, в Евросоюзе в 2018 г. была сформирована коалиция cOALition S, состоящая из организаций, финансирующих научные исследования [48]. Эта коалиция выступает за то, чтобы все результаты финансируемых государством исследований были размещены в открытом доступе. 4 сентября 2018 г. коалицией был опубликован План S, в соответствии с которым к 2020 г. все ученые, получающие деньги на исследования от государства, обязаны размещать результаты в открытом доступе [49].

План S был разработан в соответствии с десятью принципами [50].

1. Авторы сохраняют права на свои публикации, издательствам авторские права не передаются.
2. Должны быть определены критерии для журналов и площадок открытого доступа.
3. Создание платформ открытого доступа должно стимулироваться членами коалиции.
4. За публикацию материалов не должны платить исследователи, за них обязаны платить спонсоры или организации, в которых работают эти исследователи.
5. Размер платы должен быть стандартизирован и ограничен.
6. Должна быть разработана политика открытого доступа, и все обязаны действовать согласованно в рамках этой политики.
7. Для книг и монографий срок вступления в силу Плана S может быть отложен.
8. Признается важность открытых репозиториев.
9. Гибридный тип открытого доступа признается не соответствующим Плану S.
10. Члены коалиции должны контролировать соблюдение Плана S и наказывать за его нарушение.

План S имеет как сторонников, так и противников. Основными его противниками, естественно, являются издательства, которые заявляют, что реализация Плана S приведет к ухудшению качества рецензирования, уменьшению количества публикаций и сокращению распространения результатов научных исследований. Неожиданно, что против Плана S выступили некоторые научные исследователи, которые утверждают, что они поддерживают идею открытого доступа, но при этом считают, что представленный план имеет слабые места [51]. За пределами Европы отношение к Плану S тоже очень сильно различается [52].

ЗАКЛЮЧЕНИЕ

Научные исследования всегда подразумевали открытость, поскольку только те результаты, которые могут свободно обсуждаться, критиковаться, воспроизводиться другими учеными и использоваться в последующих исследованиях, могут считаться научными. Издательства, монополизировавшие рынок научной периодики, безусловно, вносят свой вклад в процесс повышения качества научных публикаций, но при этом создают барьеры для доступа к ним, что противоречит изначальным принципам распространения научных знаний.

На данном этапе мы наблюдаем некий переходный процесс, когда научное сообщество пытается выработать новые подходы к накоплению и обмену научными достижениями с использованием современных технологий [53]. При этом затрагиваются интересы множества игроков, что приводит к определенным конфликтам в данной сфере. Разрешение накопившихся проблем, по всей видимости, потребует существенного преобразования на межгосударственном уровне устоявшихся правил, политики выделения финансирования на проведение исследований, этики и культуры научных коммуникаций, законодательства в авторском праве и т.д.

СПИСОК ЛИТЕРАТУРЫ

1. Стерлигов И.А. Мерило прогресса: зачем России был нужен Web of Science и как без него дальше // Официальный сайт научно-популярного интернет-издания N+1. – URL: <https://nplus1.ru/blog/2022/05/24/no-wos-russia> (дата обращения: 28.07.2023).
2. Земсков А.И., Шрайберг Я.Л. Системы открытого доступа к информации: причины и история возникновения // Научные и технические библиотеки. – 2008. – № 4. – С. 14-29.
3. Albee B., Dingley B. U.S. Periodical Prices—2001 // American Libraries 2. – 2001. – May. – URL: <https://www.ala.org/ala/alonline/selectedarticles/periodicals01.pdf> (дата обращения: 28.07.2023)/
4. Косяков Д.В. Российская наука в открытом доступе: состояние и тенденции // Сборник тезисов докладов Международной научно-практической конференции Наука, технологии и информация в библиотеках (Libway-2019), Иркутск, 17–19 сентября 2019 г. / под общей ред. Е.Б. Артемьевой. – Иркутск: Государственная публичная научно-техническая библиотека СО РАН, 2019. – С. 107-110.
5. Глушановский А.В. Создание и эволюция централизованной системы доступа к электронным версиям научных журналов для информационной поддержки российских научных исследований. // Наука и научная информация. – 2022. – № 5(2). – С. 61-75. <https://doi.org/10.24108/2658-3143-2022-5-2-2>
6. Отчет о результатах деятельности федерального государственного бюджетного учреждения "Российский фонд фундаментальных исследований" и использовании закрепленного за ним федерального имущества за 2020 год. – URL: https://www.rfbr.ru/rffi/getimage/Отчет_о_результатах_деятельности_федерального_государственного_бюджетного_учреждения_%22Российский_фонд_фундаментальных_исследований%22_и_использовании_закрепленного_за_ним_федерального_имущества_за_2020_год.pdf?objectId=2120670&v=1690533102814 (дата обращения: 28.07.2023)
7. Отчет о результатах деятельности федерального государственного бюджетного учреждения "Российский фонд фундаментальных исследований" и использовании закрепленного за ним федерального имущества за 2021 год. – URL: https://www.rfbr.ru/rffi/getimage/Отчет_о_результатах_деятельности_федерального_государственного_бюджетного_учреждения_%22Российский_фонд_фундаментальных_исследований%22_и_использовании_закрепленного_за_ним_федерального_имущества_за_2021_год.pdf?objectId=2127682&v=1690533102814 (дата обращения: 28.07.2023)
8. Отчет о результатах деятельности федерального государственного бюджетного учреждения "Российский центр научной информации" и использовании закрепленного за ним федерального имущества за 2022 год. – URL: https://www.rfbr.ru/rffi/getimage/Отчет_о_результатах_деятельности_федерального_государственного_бюджетного_учреждения_%22Российский_центр_научной_информации%22_и_использовании_закрепленного_за_ним_федерального_имущества_за_2022_год.pdf?objectId=2132838&v=1690533102813 (дата обращения: 28.07.2023)
9. Перзановски А., Шульц Д. Конец владения: личная собственность в цифровой экономике. – Москва: Издательский дом «Дело» РАНХиГС, 2019. – 352 с.
10. Бричковский В. Инициатива открытого доступа в информационном обеспечении инновационной деятельности // Наука и инновации. – 2019. – № 12(202). – С. 76-79.
11. Земсков А.И., Шрайберг Я.Л. Системы открытого доступа к информации: причины и история возникновения // Научные и технические библиотеки. – 2008. – № 4. – С. 14-29.
12. Юрченко С.Г. "Бриллиантовый" открытый доступ в условиях самоизоляции: реализация актуальных подходов в Вестнике НМС // Вестник Научно-методического совета по природообустройству и водопользованию. – 2020. – № 17. – С. 7-16.
13. Declaration of Budapest Open Access Initiative. – URL: <https://www.budapestopenaccessinitiative.org/read/> (дата обращения: 28.07.2023)
14. Berlin Declaration. – URL: <https://openaccess.mpg.de/Berlin-Declaration> (дата обращения: 28.07.2023)
15. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. – URL: https://open-access.mpg.de/67605/berlin_declaration_engl.pdf (дата обращения: 28.07.2023).
16. Final Statement of the 16th Berlin OA conference. – URL: <https://openaccess.mpg.de/b16-final-statement?c=318911> (дата обращения: 28.07.2023).
17. Tennant JP, Waldner F, Jacques DC *et al.* The academic, economic and societal impacts of

- Open Access: an evidence-based review. – URL: <https://doi.org/10.12688/f1000research.8460.3> (дата обращения: 28.07.2023).
18. Davis P.M. Open access, readership, citations: a randomized controlled trial of scientific journal publishing // The FASEB Journal. – July 2011. – Vol. 25, Issue 7. – P. 2129-2134.
19. Москалева О.В., Акоев М.А. Прогноз развития российских журналов. Российские журналы открытого доступа // Наука и научная информация. – 2021. – Т. 4, № 1-2. – С. 33-62. DOI: 10.24108/2658-3143-2021-4-1-2-29-58.
20. Комарица В.Н. Преимущество использования открытого доступа: анализ цитирования // Научно-техническая информация. Сер. 2. – 2022. – № 7. – С. 36-41. – DOI: 10.36535/0548-0027-2022-07-4.
21. Чернова О.А. Влияние открытого доступа на научометрические показатели российских экономических журналов // Управленец. – 2022. – Т. 13, № 4. – С. 69-82. DOI: 10.29141/2218-5003-2022-13-4-6.
22. Юдина И.Г., Федотова О.А. Репозитории научных публикаций открытого доступа: история и перспективы развития // Информационное общество. – 2020. – № 6. – С. 67-79.
23. Мохначева Ю.В., Цветкова В.А. Россия в мировом массиве научных публикаций // Вестник Российской академии наук. – 2019. – Т. 89, № 8. – С. 820-830. DOI: 10.31857/S0869-5873898820-830.
24. Малахов В.А. Движение за открытый доступ к научной литературе: причины возникновения, современное состояние и перспективы развития // Управление наукой: теория и практика. – 2021. – Т. 3, № 3. – С. 118–133.
25. Наука в открытом доступе. Ученые за и против. – URL: <http://libinform.ru/read/articles/Nauka-v-otkrytom-dostupe-Uchenye-za-i-protiv/> (дата обращения: 31.07.2023).
26. Шрайберг Я.Л., Земсков А.И. Модели открытого доступа: история, виды, особенности, терминология // Научные и технические библиотеки. – 2008. – № 5. – С. 68-79.
27. Засурский И.И., Соколова Д.В., Трищенко Н.Д. Репозитории открытого доступа: функции и тенденции развития // Научные и технические библиотеки. – 2020. – № 9. – С. 121-142. DOI: 10.33186/1027-3689-2020-9-121-142.
28. Рождественская М.Ю. Репозиторий как реализация идей открытого доступа к научным публикациям: подходы к классификации // Библиосфера. – 2015. – № 2. – С. 86-94.
29. OpenDOAR Statistics: Software Platforms Overview. – URL: https://v2.sherpa.ac.uk/view/repository_visualisations/1.html (дата обращения: 31.07.2023).
30. Registry of Open Access Repositories. – URL: <http://roar.eprints.org> (дата обращения: 31.07.2023).
31. OpenDOAR. – URL: <https://v2.sherpa.ac.uk/opendoar/> (дата обращения: 31.07.2023).
32. OpenDOAR: Browse by Country and Region. – URL: https://v2.sherpa.ac.uk/view/repository_by_country/Russian_Federation.default.html (дата обращения: 31.07.2023).
33. Mulligan The effect of the Internet on researcher motivations, behaviour and attitudes // Journal Doc. – 2011. – Vol. 67, № 2. – С.290-311
34. Трищенко Н.Д. Открытый доступ к науке: анализ преимуществ и пути перехода к новой модели обмена знаниями. – Москва : Ассоциация интернет-издателей; Екатеринбург : Кабинетный учёный, 2017. – 198 с.
35. Академик РАН Алексей Хохлов: «База данных Scopus перестала быть доступной в России». – URL: <https://poisknews.ru/science-politic/akademik-ran-aleksej-hohlov-baza-dannyh-scopus-perestala-byt-dostupnoj-v-rossiidirugih-czelej/> (дата обращения: 31.07.2023).
36. OpenCitations. – URL: <http://opencitations.net/about> (дата обращения: 31.07.2023).
37. An initiative to open up citation data. – URL: <https://i4oc.org/#goals> (дата обращения: 31.07.2023).
38. Гуреев В.Н., Мазов Н.А. Возрастание роли открытых библиографических данных в условиях ограничения доступа к коммерческим информационным системам // Управление наукой: теория и практика. – 2023. – Т. 5, № 2. – С. 49-76. DOI: 10.19181/smtp.2023.5.2.4.
39. Martín-Martín A. Coverage of open citation data approaches parity with Web of Science and Scopus. – URL: <https://opencitations.wordpress.com/2021/10/27/coverage-of-open-citation-data-approaches-parity-with-web-of-science-and-scopus/> (дата обращения: 31.07.2023).
40. Открытая наука: путь в неизбежность. – URL: <http://libinform.ru/read/articles/Otkrytaya-nauka-put-v-neizbezhnost/> (дата обращения: 31.07.2023)
41. Гринберг М.Л. Подводные камни системы публикаций открытого доступа: мнения в разных странах // Научная периодика: проблемы и решения. – 2014. – № 2(20). – С. 11–20.
42. Юрченко С.Г. "Бриллиантовый" открытый доступ в условиях самоизоляции: реализация актуальных подходов в Вестнике НМС // Вестник Научно-методического совета по природообустройству и водопользованию. – 2020. – № 17. – С. 7-16.
43. Разумова И.К., Литвинова Н.Н., Шварцман М.Е., Кузнецов А.Ю. Отношение российского научного сообщества к открытому доступу: 2018 г. Анализ результатов опроса // Наука и научная информация. – 2018. – № 1(1). – С. 6-21. DOI: <https://doi.org/10.24108/2658-3143-2018-1-1-6-21>.
44. Björk B.C. Open access to scientific publications - An analysis of the barriers to change?. Information Research. – 2004. – № 9(2). – Р. 170. – URL: <http://hdl.handle.net/10227/647> (дата обращения: 31.08.2023).
45. Willinsky J. The nine flavours of open access scholarly publishing // Journal of Postgraduate Medicine. – 2003. – 49. – Р. 263–267.
46. Matsubayashi M., Kurata K., Sakai Y., Morioka T., Kato S., Morioka T., Kato S., Mine S., Ueda S. Status of open access in the biomedical field in 2005 // Journal of the Medical Library Association. – 2009. – № 97(1). – Р. 4–11.

47. Пивовар Х., Прим Д., Ларивье В. И др. Открытый доступ сегодня: широкомасштабный анализ распространенности и влияния статей открытого доступа // Наука и научная информация. – 2019. – Т. 2, № 4. – С. 228-247. DOI: 10.24108/2658-3143-2019-2-4-228-247.
48. What is cOAlition S? – URL: <https://www.coalition-s.org/about/> (дата обращения: 31.07.2023).
49. Why Plan S? – URL: <https://www.coalition-s.org/why-plan-s/> (дата обращения: 31.07.2023).
50. Plan S Principles. – URL: https://www.coalition-s.org/plan_s_principles/ (дата обращения: 31.07.2023).
51. План S: учёные поддерживают, но не хотят рисковать. – URL: <http://libinform.ru/read/articles/Plan-S-uchyonye-podderzhivayut-no-ne-hotyat-riskovat/> (дата обращения: 31.07.2023).
52. Will the world embrace Plan S, the radical proposal to mandate open access to science papers? – URL: <https://www.science.org/content/article/will-world-embrace-plan-s-radical-proposal-mandate-open-access-science-papers> (дата обращения: 31.07.2023).
53. Литвинова Н.Н., Разумова И.К. Отношение российского научного сообщества к открытому доступу: 2020. Два года спустя. Наука и научная информация. – 2020. –№3(4). С. 226-260. DOI: <https://doi.org/10.24108/2658-3143-2020-3-4-243-277>.

Материал поступил в редакцию 23.10.23.

Сведения об авторе

ШОРИН Олег Николаевич – кандидат технических наук, директор Библиотеки по естественным наукам Российской академии наук (БЕН РАН), Москва
e-mail: shorin@benran.ru

В.С. Васильева

О роли национальной киноиндустрии в экономике страны: научометрические и статистические методы оценки

Рассматривается развитие сферы киноиндустрии по отношению к экономическому развитию России. Используемые методы анализа информации основаны на научометрических и статистических показателях. Научометрические показатели включают в себя анализ роста (снижения) количества публикаций в динамике за ряд лет; публикаций российских авторов в российских и зарубежных журналах (территориальный охват); наиболее востребованных журналов, в которых авторы предпочитают публиковать результаты своих исследований; аффилиаций авторов и др. Для информационной поддержки принятия управленческих решений проведен анализ динамики статистических показателей, дополнительно проверен уровень их корреляции. Для оценки связей были сформированы матрицы парных коэффициентов корреляции, позволяющие определять силу и устойчивость связей между показателями за период 2010–2021 гг. Сделаны выводы, что почти все рассмотренные инструменты в сфере кинематографа способны воздействовать как на социальную сферу, так и на экономическое развитие страны в целом, однако эта зависимость не однонаправленная. Экономическое развитие, в свою очередь, напрямую влияет на устойчивый рост показателей, характеризующих развитие отраслей культуры России, для которых крайне важно воспроизведение и прирост высококвалифицированных кадров, развитие технологий и инноваций, в том числе, как показал анализ, посредством использования сети Интернет.

Ключевые слова: научометрические показатели, корреляционная связь, социально-экономические показатели, кино и киноиндустрия, культурная сфера, среднедушевой доход, активные пользователи сети Интернет

DOI: 10.36535/0548-0019-2024-01-4

ВВЕДЕНИЕ

За последнее десятилетие в России произошли большие изменения во всех сферах жизни: социально-экономической, культурной, научной. Культурная сфера тесно связана с социально-экономической, являясь ее неотъемлемой частью. Развитие экономики региона способствует культурному развитию: в регионах с сильной экономикой увеличиваются инвестиции в культурную сферу, строятся новые концертные залы и театры, реконструируются действующие. В то же время и развитие культурной сферы способствует экономическому росту региона, создаются новые рабочие места, улучшается благосостояние граждан, культурные мероприятия привлекают туристов, принося прибыль бизнесу региона, увеличиваются налоговые отчисления в федеральный и региональный бюджеты, регион становится более привлекателен для инвестиций.

Цель настоящего исследования – установить зависимость культурных (в том числе сферы кино и ки-

ноиндустрии) и социально-экономических показателей, оценить степень их влияния друг на друга, а также описать (интерпретировать) полученные взаимозависимости.

Исходя из того, что интерес к данной тематике характерен прежде всего для отечественных исследователей, анализ публикационной активности по теме «киноиндустрия» был выполнен на основе изучения статей российских авторов, размещенных на библиометрической платформе Scopus.

Практическая и теоретическая значимость исследования заключается в возможности использования полученных результатов органами власти и управления, ответственными за принятие решений, при формировании методологических принципов распределения затрат и доходов бюджетов российского кино во взаимоувязке с экономическими показателями и информационной поддержкой управленческих решений.

Концепция настоящего исследования базируется на двух подходах: 1) оценка публикаций по теме «Кино и киноиндустрия» на основе библиометриче-

ского анализа; 2) оценка с помощью инструментов корреляционного анализа влияния кинематографа на экономическое развитие страны.

Первый подход включает библиометрический обзор ландшафта наиболее активно цитируемых публикаций российских авторов по теме «Кино и киноиндустрия» по междисциплинарной базе данных *Scopus* за период 2017–2021 гг.

Второй подход основан на использовании инструментов корреляционного анализа основных экономических и культурных показателей за период с 2010–2021 гг. и предусматривает оценку показателей и степени их взаимного влияния (корреляции) в целом по России и по федеральным округам; выделение главных факторов (направлений) такого влияния; формирование выводов и предложений.

НАУКОМЕТРИЧЕСКИЙ АНАЛИЗ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ РОССИЙСКИХ УЧЕНЫХ В СФЕРЕ КИНЕМАТОГРАФА

Анализ публикационной активности российских исследователей в 2017–2021 гг. по теме «Кино и киноиндустрия» выполнен нами на основе данных

международной системы научного цитирования *Scopus*¹. В течение исследуемого периода в этой системе было зарегистрировано 189 публикаций российских авторов по рассматриваемой теме. Наиболее продуктивным оказался 2020 г. (47 публикаций), по сравнению с 2017 г. (27 публикаций) рост составил 42 % (рис. 1). Увеличение количества публикаций в 2020 г. можно объяснить тем, что тотальные локдауны в этот период фактически вынудили исследователей из офлайна перейти в онлайн, и тем самым появилось дополнительное время на исследовательскую работу на дому.

Пандемия коронавируса, достигшая пика распространения в 2020 г., оказала колоссальное влияние на киноиндустрию. По всему миру закрылись кинотеатры, фестивали были отменены или отложены, выпуски фильмов перенесены на неопределенный срок или ушли в онлайн.

Большинство журналов, в которых печатают свои труды российские авторы в рассматриваемом диапазоне поиска, издаются в России. На 2-м месте – публикации российских авторов, размещенные в журналах в Великобритании. Остальные страны представлены слабо (рис. 2).

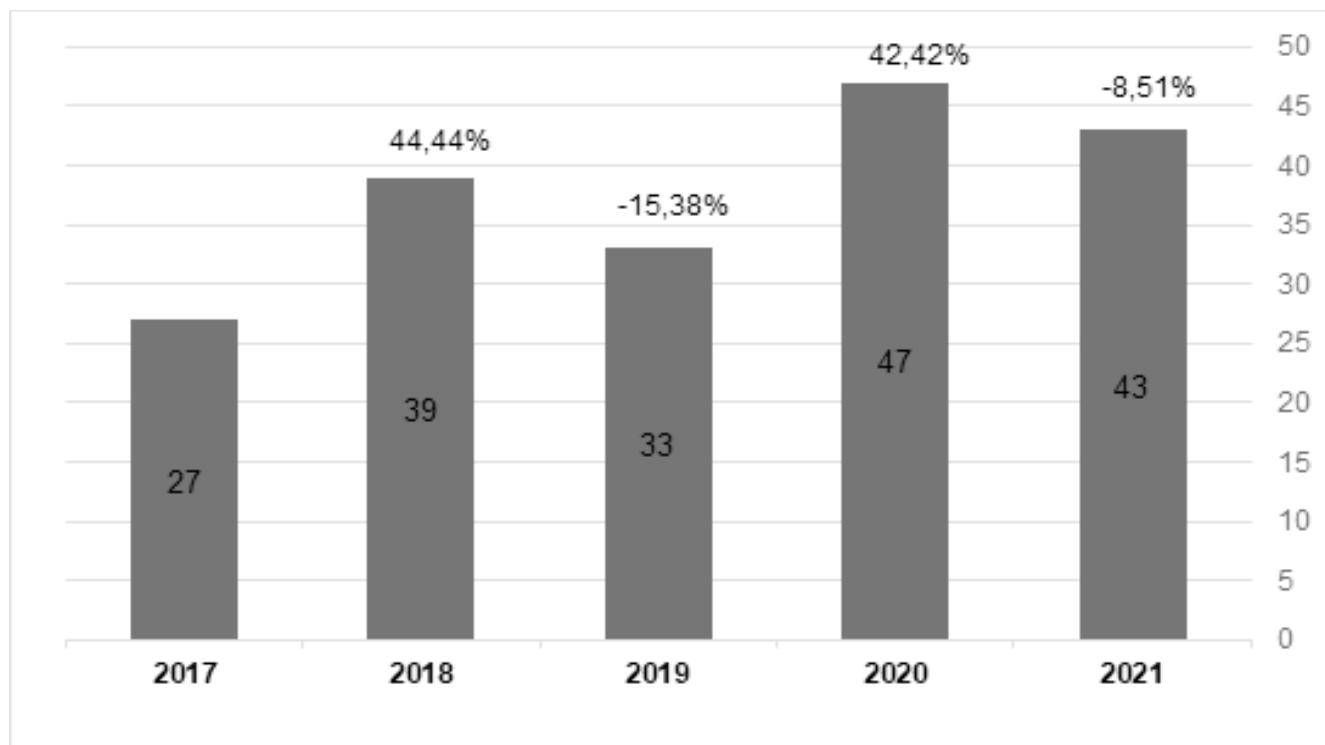


Рис. 1. Общее количество публикаций по теме «Кино и киноиндустрия», зарегистрированных в системе *Scopus* в 2017–2021 гг.

¹ Выбор платформы *Scopus* обусловлен тем, что она является одним из основных источников информации о наиболее авторитетных изданиях со всего мира из разнообразных сфер науки и техники. Анализируемый период охватывает 2017–2021 гг. В расчете не участвовал 2022 г., так как с 1 января 2023 г. отключен доступ к *Scopus* для всех российских пользователей. Доступен только режим *Scopus Preview*.

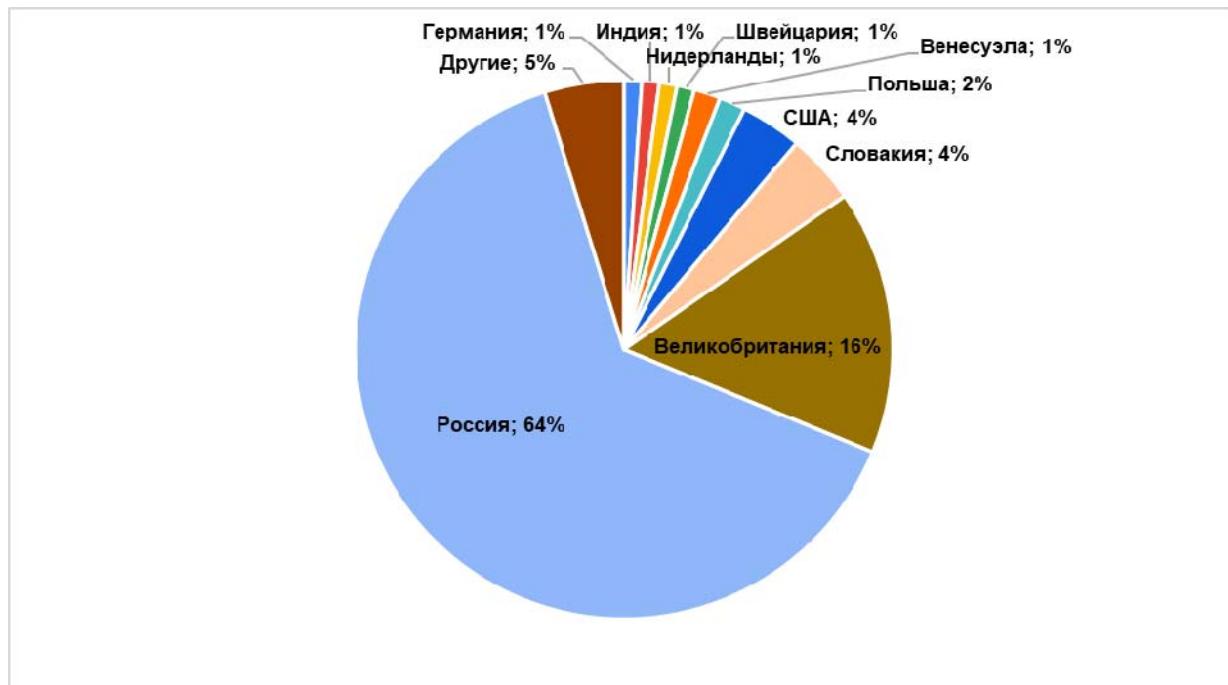


Рис. 2. Доля публикаций российских авторов по теме «Кино и киноиндустрия» в журналах по странам в 2017–2021 гг.

Преобладание публикаций российских авторов преимущественно в российских журналах, индексируемых в международных базах данных, может быть связано с геополитической обстановкой. Кроме того, публикации российских авторов в гуманитарных направлениях не так широко представлены за пределами России (рис. 3). Зарубежные (западные) журналы охотно публикуют статьи в основном по естественно-научным дисциплинам. В перспективе будет наблюдаться тенденция к снижению общего числа публикаций российских авторов в зарубежных журналах в связи с тем, что доступ российских клиентов к международным научометрическим базам данных был приостановлен в 2022 г. Однако они могут размещать свои статьи в журналах, индексируемых *Russian Science Citation Index (RSCI)*², и в журналах стран, дружественных России.

Сегодня российское руководство нацелено на активное внедрение собственной системы оценки эффективности научных исследований. В сложившихся условиях адаптации России к новым вызовам государственные институты должны обратить внимание на российскую базу журналов *RSCI* как альтернативу западным и предусмотреть финансируемую государством программу ее развития (*Rg.ru*, 2022). В декабре 2022 г. в сфере отечественной науки произошли значительные изменения: все научные журналы списка ВАК были поделены на 3 большие группы – квартили K1, K2 и K3 [1]. Теперь статьи, опубликованные в журналах квартиля K1, будут приравнены по статусу

к статьям, опубликованным в журналах, внесенным ранее в международные базы данных.

Наибольшее количество статей российских авторов по рассматриваемой теме за период 2017–2021 гг. было опубликовано в журнале «Исследования российского и советского кинематографа» (Великобритания), специализирующегося на публикации статей в области искусства и СМИ (18 ед.). В Топ-10 также вошли российские и зарубежные журналы, однако они представлены более слабо – 5–7 публикаций (рис. 4).

Интересная закономерность отмечена при анализе научных публикаций в их привязке к организациям в соответствии с аффилиацией авторов. Наибольшее количество статей по теме исследования наблюдается у организаций, для которых данная деятельность не является профильной. Публикации представлены рейтинговыми федеральными университетами. Однако в топ-5 не попали профильные институты, для которых область театра и кино была основной (рис. 5).

Наибольшая доля публикаций – статьи, в меньшей степени – обзоры (рис. 6). При этом все остальные типы публикаций – труды конференций, письма, заметки, книги, главы из книг и т. д. – оказываются вне поля зрения. Это связано с требованием учета публикаций только определенных типов – исследовательских статей и обзоров – при использовании научометрических показателей [2].

Публикации представлены в основном на английском и русском языках; лишь одна статья исследователя НИУ ВШЭ была издана на французском языке в журнале *«Revue des études slaves»* (Журнал славистики) (рис. 7).

Начиная с 2018 г. преобладают публикации на русском языке, достигнув максимального значения в 2021 г. (рис. 8).

² *RSCI* – это так называемая российская полка журналов на платформе *Web of Science*, к которым относят авторитетные российские периодические издания, индексируемые в международных базах данных. База *Web of Science* стала недоступной в России в 2022 г.



Рис. 3. Количество статей российских авторов в российских и зарубежных журналах по теме «Кино и киноиндустрия» в 2017–2021 гг.



Рис. 4. Топ-10 журналов, в которых опубликованы результаты исследований российских авторов по теме «Кино и киноиндустрия» в 2017–2021 гг.

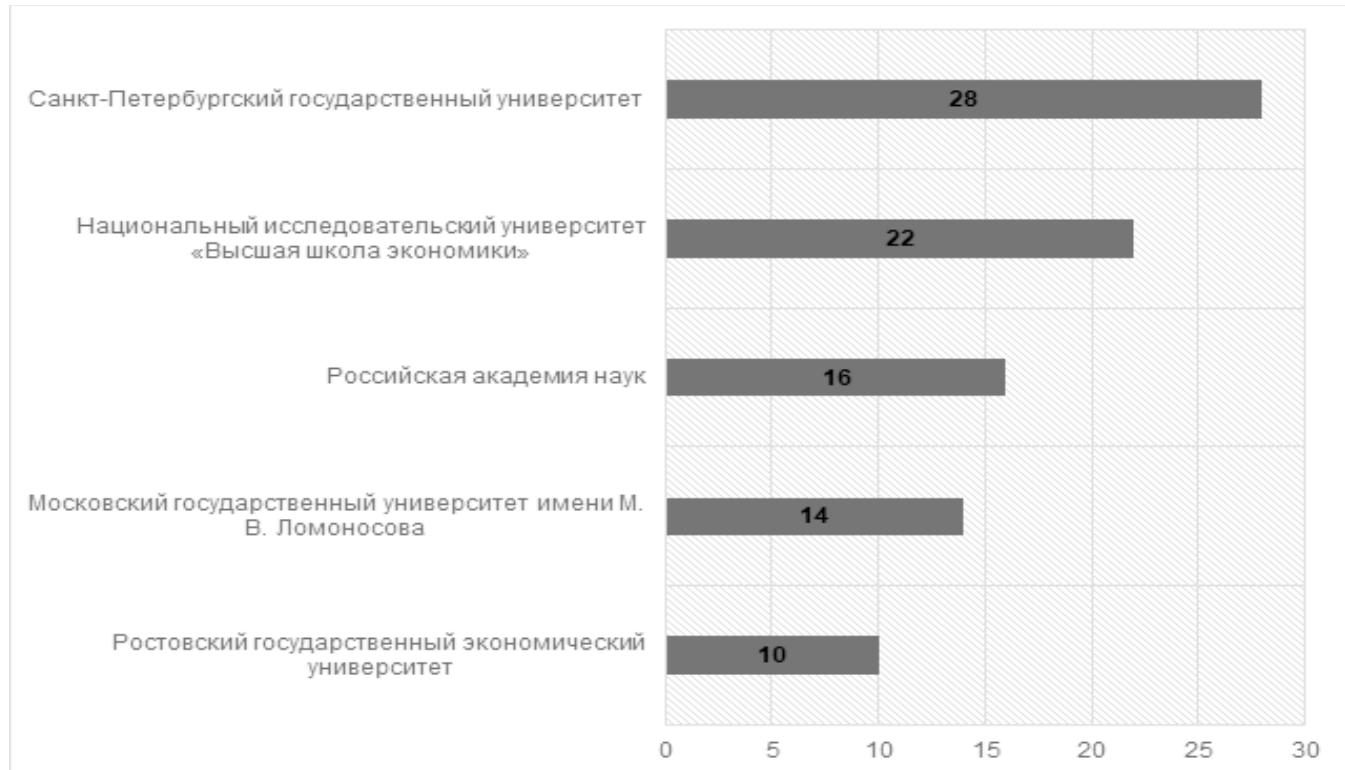


Рис. 5. Топ-5 организаций, публикующих статьи по теме «Кино и киноиндустрия» в 2017–2021 гг.

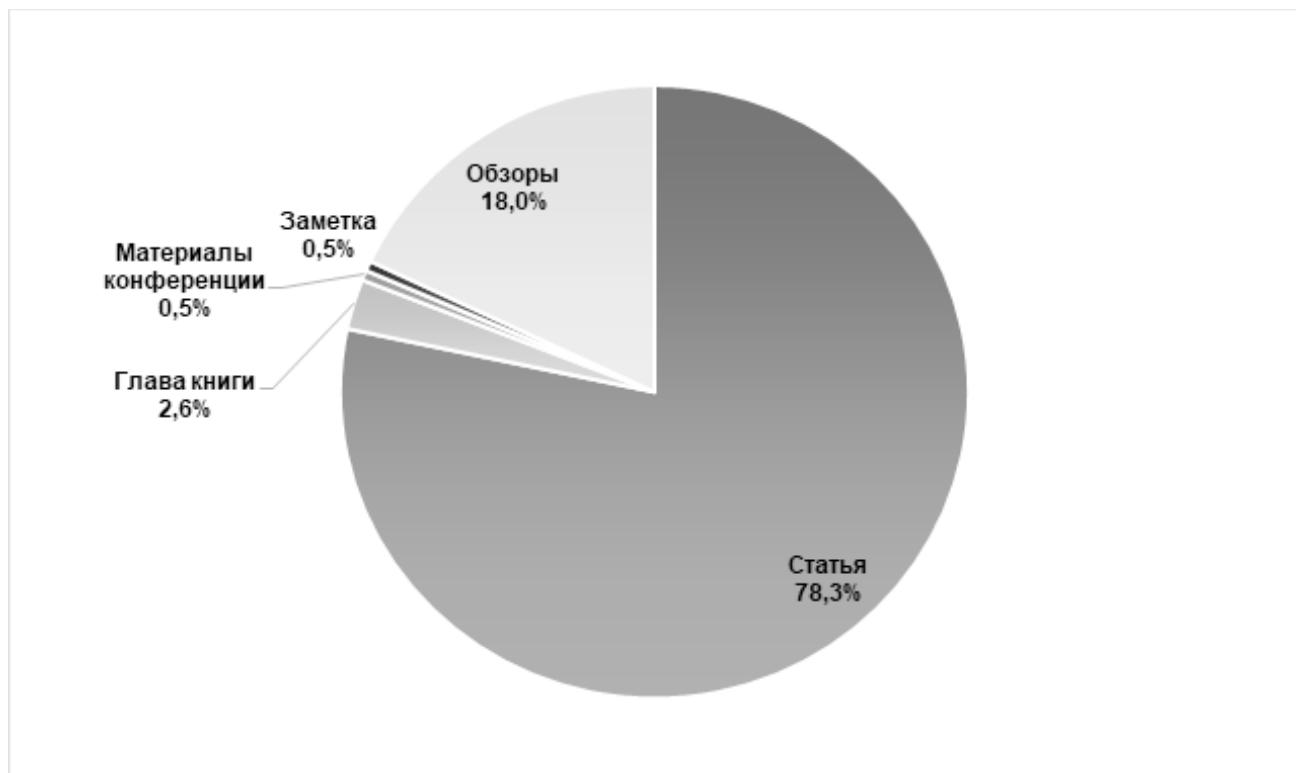


Рис. 6. Тип публикаций по теме «Кино и киноиндустрия» в 2017–2021 г.

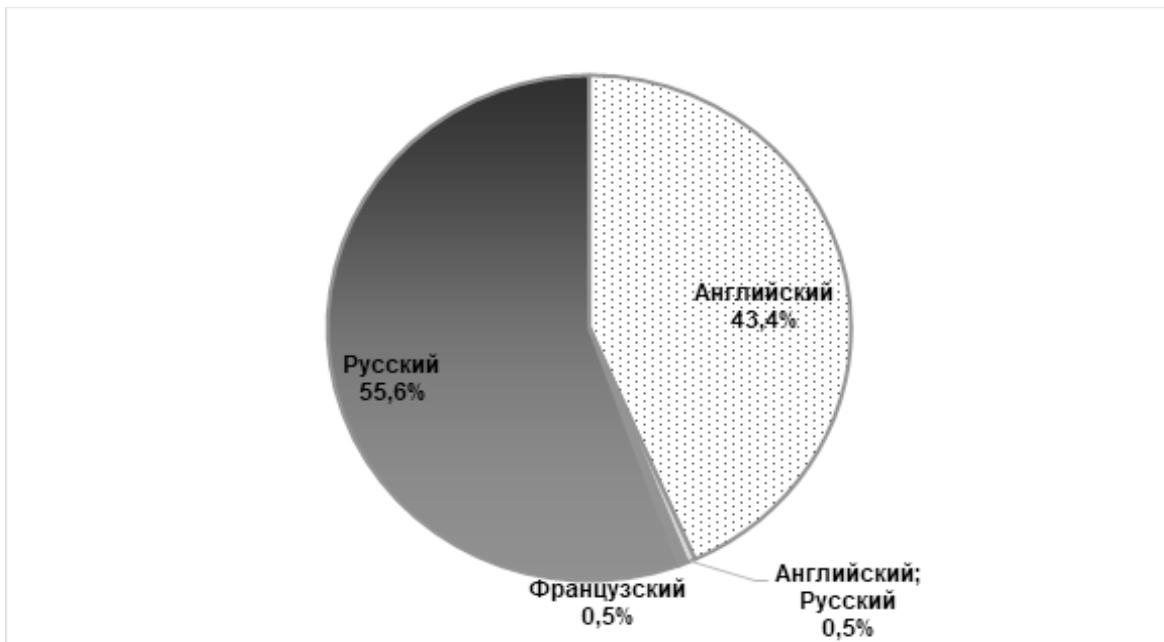


Рис. 7. Языки публикаций по теме «Кино и киноиндустрия» в 2017–2021 гг.

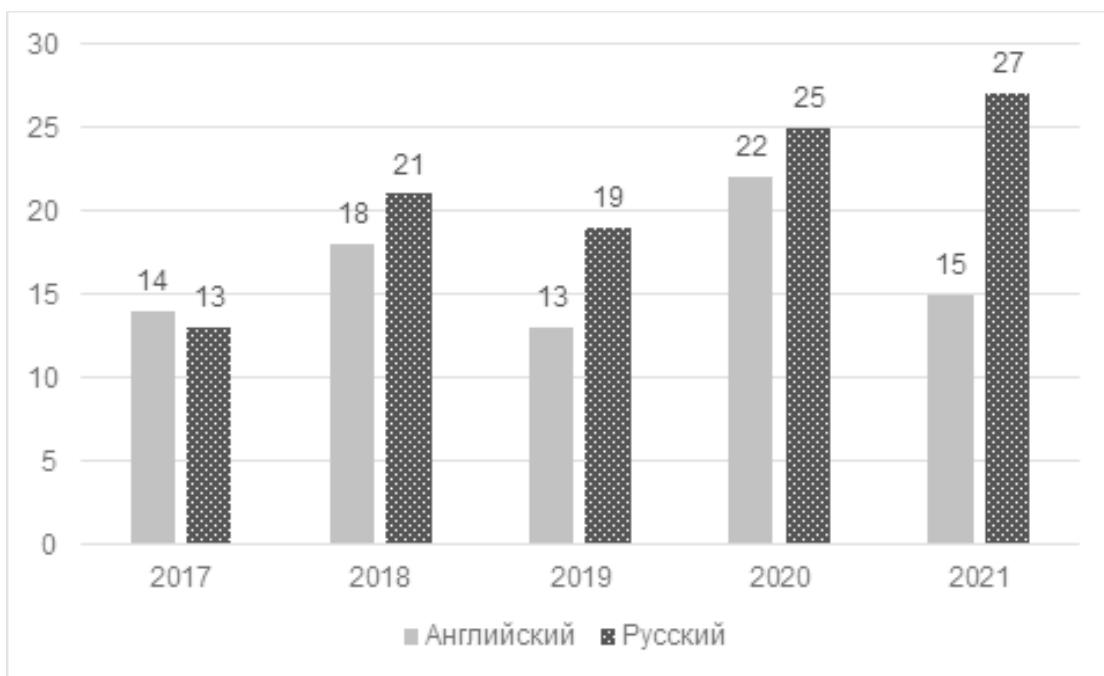


Рис. 8. Количество публикаций на русском и английском языках по теме «Кино и киноиндустрия» в 2017–2021 гг.

Таким образом, в целом можно отметить общую тенденцию роста числа статей (хотя и нестабильного) по тематике «Кино и киноиндустрия» в журналах, индексируемых в международных базах данных, который, однако, достигнут преимущественно за счет публикаций в российских журналах авторов с аффилиацией федеральных непрофильных университетов России.

Следует отметить, что, поскольку кинематограф выполняет социальные функции в обществе, большое значение имеет представленность российских теоретических и практических разработок по данной тематике как на отечественной, так и международной арене.

На наш взгляд, необходимо активно содействовать развитию публикационной активности среди профессионалов, имеющих опыт, среди молодежи, обучающейся в театральных вузах и институтах кино. Многие из них имеют порой нестандартные и заслуживающие внимания мнения и подходы по становлению киноиндустрии как способу развития культуры социума, взгляды на киноиндустрию не только как на сферу культуры, но и как на сферу бизнеса, что в определенной степени может стать триггером для выхода России на новый уровень в киноискусстве.

ПОСТАНОВКА ПРОБЛЕМЫ, МЕТОДИКА ИССЛЕДОВАНИЙ И ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Проблема влияния кинематографа на экономику, на наш взгляд, недостаточно изучена. В связи с этим мы осуществили сбор необходимой статистической информации и проанализировали взаимосвязи показателей, характеризующих кинематограф, с показателями социально-экономического развития России. Такие связи выявляют инструменты в сфере кинематографа, способные воздействовать на социальную сферу или экономическое развитие, и определяют, насколько такое воздействие сильное и масштабное. Это может иметь не однородный характер, поэтому следует иметь в виду обратное воздействие происходящих изменений экономических показателей и их влияние на сферу киноиндустрии.

В нашем исследовании использован библиометрический подход, который характеризует уровень публикационной активности российских ученых на основе анализа их публикаций, размещенных в международной базе данных *Scopus*, за период с 2017–2021 гг. Наряду с общим количеством публикаций по теме «Кино и киноиндустрия» анализ позволил выявить: долю таких публикаций в журналах по странам; наиболее популярные журналы, в которых опубликованы результаты исследований российских авторов; российские организации, активно публикующие статьи по теме исследования; типы и языки публикаций.

Основной метод оценки взаимосвязей социально-экономического и культурного развития регионов – это корреляционно-регрессионный анализ. Он позволяет выявить степень зависимости и плотности связи между показателями – индикаторами эффективности развития сферы киноиндустрии и социально-экономическими показателями. Результаты анализа помогут установить степень влияния сферы киноиндустрии на социально-экономическое развитие страны и выделить главные факторы (направления) такого влияния, а также прогнозировать и планировать становление основных направлений и, соответственно, принимать решения.

Как уже было отмечено, для изучения взаимосвязей социально-экономического и культурного развития регионов мы использовали корреляционный анализ, позволяющий устанавливать взаимосвязи между показателями, по которым выявлены наиболее устойчивые корреляционные связи (табл. 1).

Для анализа тенденций изменения показателей был выбран период 2010–2021 гг. Выявлены взаимосвязи показателей регионов РФ по федеральным округам, а также в целом по Российской Федерации. При этом 2020 и 2021 гг. требуют отдельного рассмотрения в связи с тем, что они характеризуются аномально низкими показателями в культурной сфере, что искажает общую картину. Это вызвано тем, что во время пандемии в 2020 г. были закрыты для посещения кинотеатры, театры, многие мероприятия были отменены. В 2021 г. действовали ограничения по заполняемости залов в связи с ковидными ограничениями.

Таблица 1

Показатели социально-экономического и культурного развития регионов, по которым установлены корреляционные связи

№ п/п	Показатели, отражающие развитие отраслей культуры региона (в т. ч. сферы киноиндустрии)	Показатели социально-экономического развития региона
1	Сумма валовых сборов от платных показов фильмов, руб. (EAIS, 2021) [3]	Валовый региональный продукт (ВРП), млн руб. (Rosstat, 2021) [6]
2	Среднее число зрителей на мероприятиях театров (в расчете на 1 тыс. чел. населения) (EMIS, 2021) [4]	Инвестиции в основной капитал, млн руб. (Rosstat, 2021)
3	Число зрителей, присутствующих на мероприятиях, проводимых театрами Минкультуры России, чел. (EMIS, 2021) [4]	Инвестиции в Российскую Федерацию, млн. руб. (Rosstat, 2021)
4	Доля российских фильмов в прокате, % (ProfiCinema, 2021) [5]	Численность занятых в возрасте 15 лет и старше по субъектам Российской Федерации, тыс. чел. (Rosstat, 2021)
5		Среднедушевые денежные доходы (в месяц), руб. (Rosstat, 2021)
6		Среднемесячная номинальная начисленная заработка плата работников по полному кругу организаций в целом по экономике по субъектам Российской Федерации, руб. (Rosstat, 2021)
7		Стоймость основных фондов (млн руб.) (Rosstat, 2021)
8		Численность активных абонентов фиксированного широкополосного доступа к сети Интернет на 100 человек, (ед.) (Rosstat, 2021)
9		Численность активных абонентов мобильного широкополосного доступа к сети Интернет на 100 человек, (ед.) (Rosstat, 2021)

С помощью пакета для анализа данных в *MS Excel* мы построили матрицы парных коэффициентов корреляции социально-экономических показателей и показателей, отражающих культурное развитие по Российской Федерации и по федеральным округам РФ. Положительные корреляционные связи выявлены между четырьмя культурными показателями (в том числе показателями в сфере кинопроизводства) и семью показателями социально-экономического развития. Наиболее сильные взаимосвязи проявились между показателем «численность активных абонентов фиксированного широкополосного доступа к сети Интернет» и такими показателями, как сумма валовых сборов от платных показов фильмов, доля отечественных фильмов в прокате и количество зрителей на мероприятиях театров, т. е. практически со всеми показателями, которые входили в изучаемую выборку.

Высокие корреляционные связи со среднедушевыми денежными доходами населения (коэффициент корреляции – 0,91), численностью активных абонентов фиксированного широкополосного доступа к сети Интернет (0,89), валовым региональным продуктом (0,88) имеет **сумма валовых сборов от платных показов фильмов**, что подтверждает сильное взаимное влияние социально-экономического и культурного развития страны.

График на рис. 9 демонстрирует, что показатели «среднедушевые доходы населения» и «сумма вало-

вых сборов от платного показа фильмов» изменяются односторонне, исключение составляют 2020 и 2021 гг. (не учитывались при расчете коэффициентов корреляции). Следует отметить снижение валовых сборов от показов фильмов в 2018 г., при этом среднедушевой доход населения увеличивается. Это можно объяснить несколькими факторами: чемпионат мира по футболу летом 2018 г., более привлекательный для части зрителей, и трагедия в кинотеатре ТРЦ «Зимняя вишня» в г. Кемерово, после которой посещаемость кинотеатров в определенной степени снизилась.

Сумма валовых сборов от платных показов фильмов имеет также высокую положительную корреляцию с численностью активных абонентов сети Интернет (коэффициент корреляции – 0,89), и показатели растут односторонне (рис. 10). Несмотря на то, что увеличивается количество онлайн-кинотеатров, можно было бы ожидать оттока зрителей из кинозалов, но это не происходит. Примечательно, что развитие сервисов покупки билетов, доступность информации о премьерах, сюжете, публикация рейтингов фильмов способствуют увеличению кассовых сборов в кинозалах. Высокие взаимосвязи эти показатели имеют как в целом по России, так и по федеральным округам. Исключение составляют Южный и Дальневосточный федеральные округа, в которых не удалось установить значимые корреляционные связи.

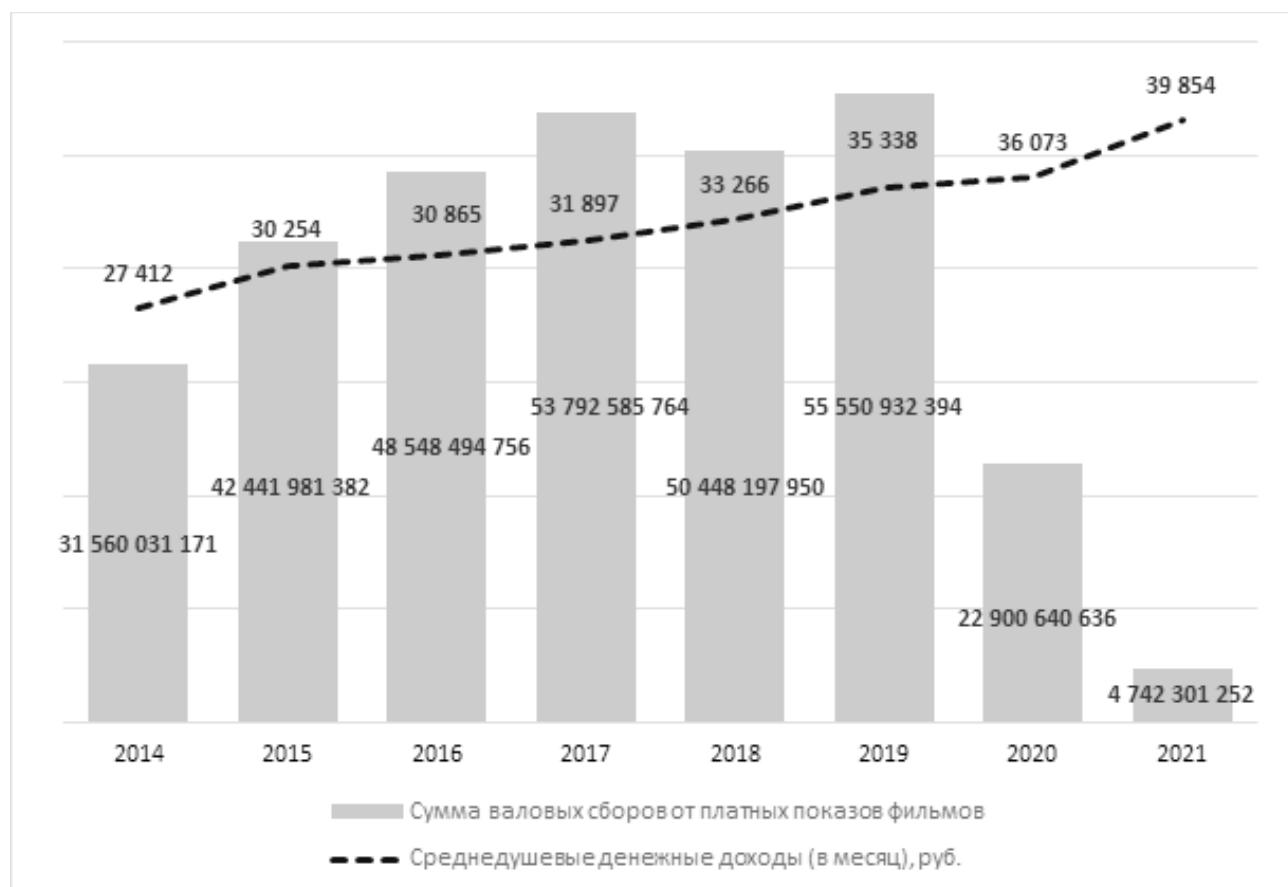


Рис. 9. Динамика суммы валовых сборов от платных показов фильмов и среднедушевого дохода в России в 2014–2021 гг.



Рис. 10. Динамика суммы валовых сборов от платных показов фильмов и численности активных российских абонентов фиксированного широкополосного доступа к сети Интернет на 100 чел. населения в 2014 – 2021 гг.

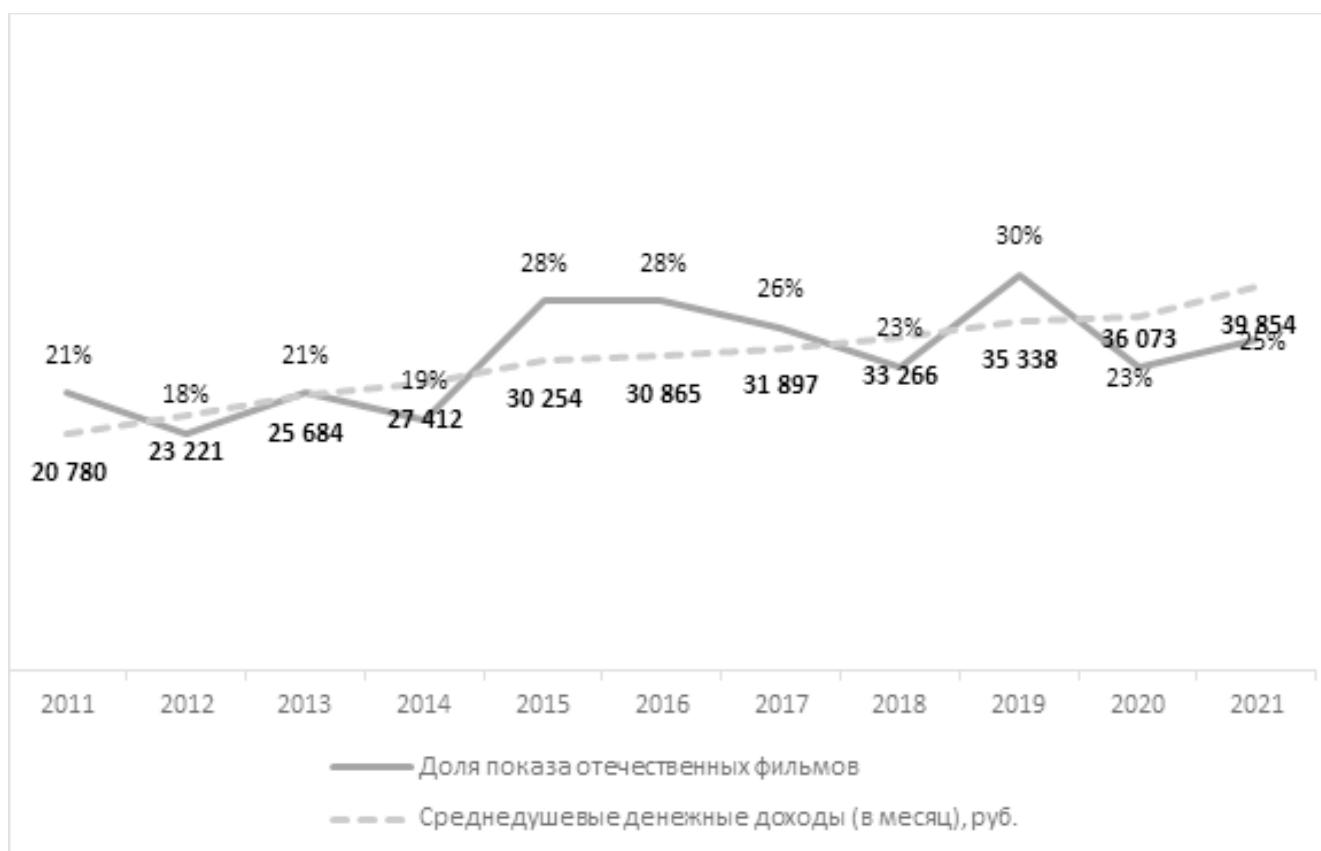


Рис. 11. Взаимосвязь доли отечественных фильмов в прокате и среднедушевого дохода в 2011 – 2021 гг.

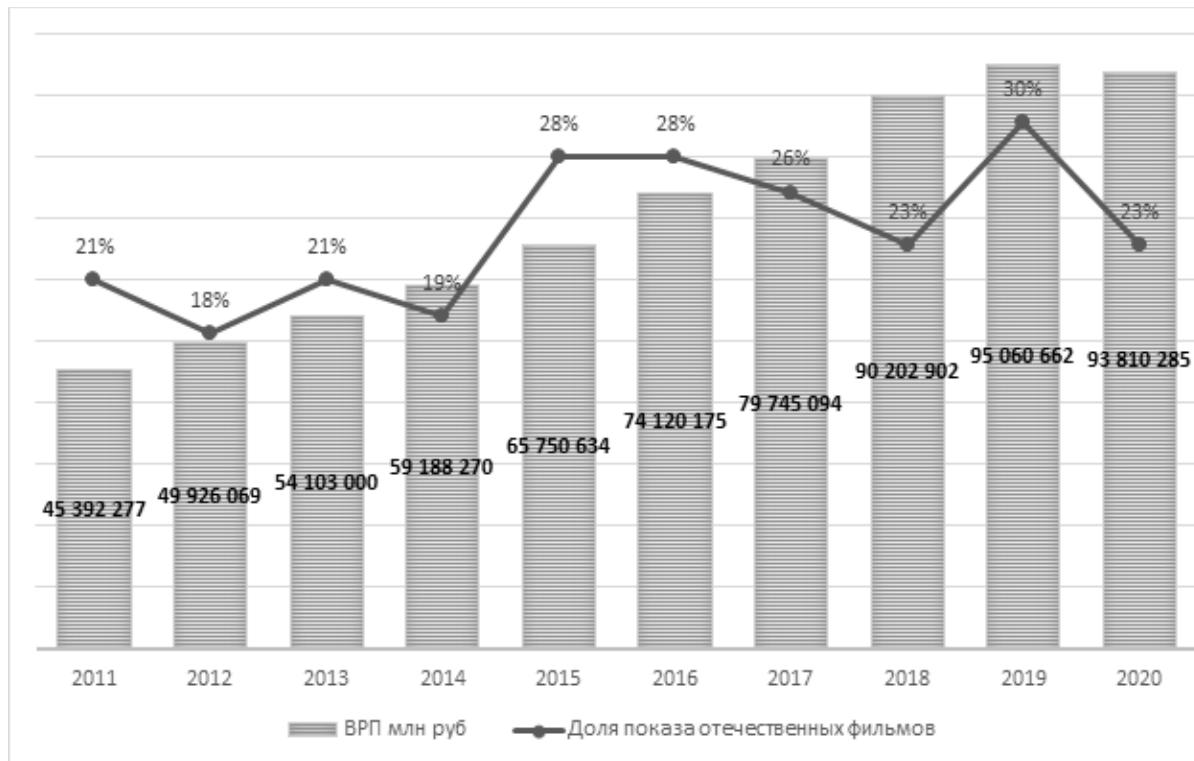


Рис. 12. Взаимосвязь доли отечественных фильмов в прокате и валового регионального продукта в 2011 – 2020 гг.
(по данным Росстат актуальные сведения по ВРП представлены за 2020 г.)



Рис. 13. Динамика числа зрителей, присутствующих на мероприятиях, проводимых театрами Минкультуры России, и численности активных абонентов фиксированного широкополосного доступа к сети Интернет на 100 чел. населения в 2011 – 2021 гг.

Доля отечественных фильмов в прокате имеет сильные корреляционные связи со среднедушевым доходом, численностью активных абонентов широкополосного фиксированного и мобильного Интернета. Следует отметить, что оценить эти связи возможно только в целом по России, так как статистические данные по доле отечественных фильмов в прокате представлены только в целом по стране.

На рис. 11 показано, что наблюдается рост тренда по показателям «доля отечественных фильмов в прокате» и «среднедушевой доход». С высокой степенью вероятности можно утверждать, что существует достаточно высокий уровень взаимосвязи между увеличением доли фильмов отечественного производства и ростом среднедушевого дохода населения. При расширении производства отечественных фильмов создаются новые рабочие места, что улучшает благосостояние работников, занятых в данной отрасли, и оказывает влияние на среднедушевой доход в целом.

Если рассмотреть взаимосвязь показателя «доля отечественных фильмов в прокате» с основным показателем, отражающим экономическое развитие страны и регионов – валовым региональным продуктом (ВРП) (рис. 12), то видно, что рост ВРП напрямую влияет на развитие кинематографической отрасли в стране: экономический рост способствует росту инвестиций в кинематограф со стороны государства и частных компаний. В то же время кинематограф влияет на изменение ВРП. При увеличении доли российских фильмов создаются рабочие места, в кинопрокате появляется больше различных новинок, что привлекает зрителей в кинотеатры и увеличивает кассовые сборы. Это отражается и на развитии смежных отраслей экономики страны (общепит, торговля и др.), так как посещение зрителями кинозалов сопровождается посещением буфетов.

Похожие тенденции наблюдаются и с посещением театров. Два показателя **«среднее число зрителей на мероприятиях театров»** и **«число зрителей, присутствующих на мероприятиях»** имеют высокие положительные корреляционные связи с такими экономическими показателями, как «инвестиции в основной капитал», «среднедушевые денежные доходы», «стоимость основных фондов», «ВРП», «численность активных абонентов фиксированного и мобильного широкополосного доступа к сети Интернет», что свидетельствует о взаимном влиянии социально-экономических показателей и показателей, отражающих развитие отраслей культуры России (рис. 13).

Наименьшую связь с выбранными нами показателями культурного развития страны демонстрирует такой показатель, как инвестиции в Российскую Федерацию. Значения коэффициентов корреляций не превышают 0,6, а для числа зрителей театров вообще находятся близко к нулю, уходя в область отрицательных значений. Это свидетельствует о том, что сфера культуры не поддерживается иностранными инвестициями и развивается только и исключительно за счет внутренних ресурсов.

Исходя из того, что наибольшую плотность имеют корреляционные связи всех четырех показателей

сферы культуры с численностью активных абонентов сети Интернет, а такими абонентами являются, как правило, работающие активные граждане [7, 8], можно сделать вывод, что сфера культуры будет наиболее активно развиваться в тех регионах, где доля таких пользователей сети Интернет в общей численности населения выше. Этот вывод подтверждает высокая плотность связи показателей сферы культуры с показателями среднедушевых доходов граждан.

Представляет интерес связь с показателями сферы культуры таких показателей экономического развития, как стоимость основных фондов и инвестиции в уставной капитал. Сфера культуры не является капиталоемкой, стоимость ее основных фондов (зданий театров, музеев, кинотеатров и проч.) не может сравняться с промышленными предприятиями, и тем не менее наблюдается положительная корреляция невысокой плотности всех 4-х рассматриваемых показателей сферы культуры с обоими показателями. Причем, именно инвестиции имеют наибольшую плотность с показателями, связанными с театральной деятельностью (зрители на мероприятиях театров), что может указывать на несколько большую активность в области строительства и реконструкции театров, чем кинотеатров, или на их более высокую капиталоемкость. При наличии статистических данных о стоимости основных фондов сферы культуры можно было бы делать более увереные выводы о развитии в этой области.

ЗАКЛЮЧЕНИЕ

В современном мире все глобальные процессы, включая экономические, обусловлены культурой. Некоторые исследователи видят в культурных ценностях стратегический ресурс для эффективного движения экономической деятельности [9].

Сфера духовной культуры: искусство, кино и театр, с одной стороны, и экономика – с другой, неразрывно связаны между собой. Соответственно, существует реальная возможность отследить культурные предпосылки для области киноиндустрии, которые могут способствовать экономическому успеху. Их выявление позволит использовать эти знания не только для решения вопросов, связанных с управлением ресурсами в этих сферах, но и в целом – для повышения конкурентоспособности российской экономики. Актуальность нашего исследования обусловлена возрастающим взаимным влиянием показателей социокультурного развития (прежде всего сферы киноиндустрии) и основных показателей, характеризующих экономическое развитие страны.

Таким образом, на основе библиометрического анализа публикаций российских исследователей в базе данных *Scopus* по теме «Кино и киноиндустрия» за период с 2017–2021 гг. было выявлено:

- наблюдается тенденция роста числа публикаций российских авторов по теме «Кино и киноиндустрия», особенно заметная начиная с 2020 г.;
- большинство журналов, в которых печатают свои труды российские авторы, издаются в России;
- наиболее востребованные журналы: «Исследования российского и советского кинематографа», «Вопросы истории», «Праксема», «Диалог со временем»;

- наибольшее количество публикаций выявлено у организаций, для которых данная деятельность не является основной профильной. В тройку лидеров вошли СПбГУ, НИУ ВШЭ, РАН;
- начиная с 2018 г. в публикациях преобладает русский язык, достигнув максимального разрыва с другими языками в 2021 г.

Корреляционный анализ основных экономических и культурных показателей позволил выявить:

- показатель «сумма валовых сборов от платных показов фильмов» коррелирует с показателями «среднедушевые доходы населения» и «численность активных абонентов сети Интернет» (на всей территории России, за исключением Южного и Дальневосточного федеральных округов);
- показатель «доля отечественных фильмов в прокате» коррелирует с показателями «среднедушевой доход», «численность активных абонентов широкополосного фиксированного и мобильного интернета» и «ВРП»;
- показатели «среднее число зрителей на мероприятиях театров» и «число зрителей, присутствующих на мероприятиях» имеют положительные корреляционные связи с большинством рассматриваемых в рамках нашего исследования экономических показателей;
- все четыре рассматриваемых показателя сферы культуры имеют тесные корреляционные связи с численностью активных абонентов сети Интернет.

Проведенный анализ позволил сделать вывод, что рассмотренные нами инструменты в сфере кинематографа способны воздействовать как на социальную сферу, так и на экономическое развитие страны в целом.

Большинство рассмотренных нами показателей имеют связь с численностью активных абонентов сети Интернет, это является подтверждением необходимости для тех, кто развивает кинобизнес, быть в курсе основных тенденций в технологиях продвижения контента через сеть Интернет.

Пользовательские привычки в Интернете меняются — предпочтения и ожидания интернет-аудитории изменяются под влиянием культурных традиций разных частей света. Социальные сети упрощают процесс взаимодействия с аудиторией [10]. Большое значение в данном контексте имеет развитие профессиональных социальных сетей, в частности, социальных сетей киноактеров и других профессиональных деятелей сферы киноискусства, а также кинофорумов.

Кинематограф в целом — это мощное культурно-пропагандистское средство. Когда государство находится в кризисе, первое, что может ему помочь, — это культурное возрождение. Культурный прогресс тесно

взаимосвязан с прогрессом экономическим и социальным, что наглядно продемонстрировали результаты настоящего исследования.

СПИСОК ЛИТЕРАТУРЫ

1. Сибак.инфо. Журналы ВАК разделили на квартили (к1, к2, к3). – URL: <https://sibac.info/blog/zhurnaly-vak-razdelili-na-kvartili-k1-k2-k3>
2. Можначева Ю. В. Классификация публикаций по типам документов в Web of Science и Scopus: сходства, различия и их значение при анализе публикационной активности // Управление наукой: теория и практика. – 2022. – Vol. 4(3). – С. 154–170. DOI: <https://doi.org/10.19181/smtp.2022.4.3.11>.
3. Единая межведомственная информационно – статистическая система (ЕМИСС). – URL: <https://www.fedstat.ru/indicators/>
4. Единая федеральная автоматизированная информационная система сведений о показах фильмов в кинозалах (ЕАИС). – URL: <https://ekinobilet.fond-kino.ru/>
5. Информационный портал для профессионалов кинобизнеса ProfiCinema. – URL: <https://www.proficinema.com/questions-problems/articles/detail.php?ID=347703>
6. Федеральная служба государственной статистики (Росстат). – URL: <https://rosstat.gov.ru/>
7. Заморкин А. А. Исследование портрета активного пользователя сети интернет // Вестник НГУЭУ. – 2013. – № 1. – С. 181–185. – URL: <https://cyberleninka.ru/article/n/issledovanie-portreta-aktivnogo-polzovatelya-seti-internet/viewer>
8. Экспертный центр Digital Russia. Рейтинг регионов России по активности пользования Интернетом. – URL: <https://d-russia.ru/reiting-regionov-rossii-po-aktivnosti-polzovaniya-internetom.html>
9. Куприна Т.В. Влияние культурных показателей на экономическую деятельность международных компаний // Экономика региона. – 2021. – Vol. 17(2). – С. 593–606. – URL: [10.17059/ekon.reg.2021-2-17](https://doi.org/10.17059/ekon.reg.2021-2-17). – EDN WTESQJ
10. WebCanape. – URL: <https://www.web-canape.ru/business/internet-2020-globalnaya-statistika-i-trendy/>

Материал поступил в редакцию 03.10.23.

Сведения об авторе

ВАСИЛЬЕВА Владлена Сергеевна – Институт кино, Национальный исследовательский университет «Высшая школа экономики», Москва
e-mail: vasilyeva128@mail.ru

СПРАВОЧНО-ИНФОРМАЦИОННЫЙ РАЗДЕЛ

УДК 002.6:004.9«Партархив»

А.Б. Антопольский, К.В. Вигурский, Ю.Г. Коргунюк

О модернизации информационной системы «Партархив»

Описываются принципы и основные параметры начатой в ИНИОН РАН модернизации информационной системы «Партархив», которая относится к смешанному документально-энциклопедическому типу. Рассматривается структура системы и правила представления основных информационных объектов: персона, организация, текстовые и мультимедийные документы, а также связи между ними. Подробно показана технология подготовки информации, включая управление этой технологией. Представлены пользовательские и административные функции, в том числе возможности навигации и поиска.

Ключевые слова: информационная систем «Партархив», модернизация, информационные объекты, атрибуты, функциональность, технология

DOI: 10.36535/0548-0019-2024-01-5

ВВЕДЕНИЕ

В ИНИОН РАН принято решение провести модернизацию и затем начать эксплуатацию информационной системы (ИС) «Партархив» (<http://www.partinform.ru/pa98/>). Эта система, ранее функционировавшая при фонде «Информатика для демократии» (Фонд ИНДЕМ), использовалась главным образом для политологических исследований [1–3]. Однако она имела ряд технологических ограничений и недостаточно применялась для информационного обслуживания ученых-политологов. За 30 лет эксплуатации в системе был накоплен громадный (около 800 тыс.) массив документов (опубликованных, сетевых, архивных и др.), характеризующих политическую жизнь России с позднесоветского периода по настоящее время. Сохранение этого массива и организация широкого общественного доступа к нему представляется важной и необходимой задачей. В то же время для этого массива необходима верификация, которая может быть весьма трудозатратной.

По сравнению с другими ресурсами преимущество ИС «Партархив» заключается в возможности доступа к материалам, которые очень трудно или вообще невозможно найти в других местах (в том числе по причине закрытия сайтов), и их сортировки по разным параметрам (организации, персоны, политические направления, даты, регионы и пр.).

Потенциальная аудитория ИС «Партархив» – это исследователи в области социальных наук (прежде

всего политологи и историки), студенты, аспиранты и соискатели ученых степеней соответствующих специальностей, а также все интересующиеся изучением общественно-политической истории России на уровне не только литературы, но и первичных источников. Эта Система должна включать и обычную политологическую научную литературу, в том числе из фондов ИНИОНа, которой ранее было явно недостаточно.

Для полнотекстовых информационных систем, таких как электронные библиотеки, большое значение имеют правовые требования, прежде всего, соблюдение авторского права. Легитимность включения информации в ИС «Партархив» гарантируется прежде всего использованием исключительно открытых источников и исполнением требований, предъявляемых к публикуемой информации обладателями авторских прав. Саму ИС «Партархив» предполагается разместить в Интернете под лицензией *Creative Commons Attribution (CC-BY)* и она должна быть доступна для скачивания и некоммерческого применения со ссылкой на источник.

Настоящая статья подготовлена на основе технического задания к ИС «Партархив», в ходе разработки которого были согласованы основные параметры её модернизации и использованы технические решения, реализованные в информационной системе «Сравнительная поэтика и сравнительное литературоведение (СПСЛ)» (<https://cpcl.info/>) [4, 5].

ИНФОРМАЦИОННАЯ СИСТЕМА «ПАРТАРХИВ»

Общие требования

Данные для Системы отбирают администратор и члены его команды в соответствии с критерием тематической направленности: документы должны иметь отношение к публичной стороне политического процесса, прежде всего к общенациональной идеино-политической дискуссии, с акцентом на деятельность общественно-политических организаций.

Верификация данных выполняется прежде всего по источнику их происхождения. Предпочтение отдается документам – артефактам деятельности организаций и персон, которые публикуются на принадлежащих им ресурсах. Для материалов СМИ реализуется принцип нескольких независимых друг от друга источников.

В ИС «Партархив» включается информация на русском языке.

Основные информационные объекты и их атрибуты

ИС «Партархив» – это смешанная документально-энциклопедическая система, включающая пять основных видов информационных объектов:

- персона,
- организация,
- документ текстовый,
- документ мультимедийный (изображение, видео, аудио),
- словари.

Объекты – это набор атрибутов. Каждый объект имеет атрибут «короткое имя», предназначенный для его представления в различных списках, в том числе в результатах поиска, для именования окон, содержащих сам объект, для средств навигации и других целей.

Заранее определенный список допустимых значений атрибутов представляет собой объект «словари».

Объекты «персона» и «организация» оформляются в виде статей, содержащих заголовок, набор атрибутов (имя атрибута и его значение) и возможно фотогалерею — набор фотографий для данного объекта. Атрибуты с отсутствующими значениями (пустые) пользователю не показываются. При этих объектах приводится (в виде коротких имен) список связанных с ними документов. Для объекта «персона» атрибутами могут быть биографические и контактные данные, сведения о политическом секторе и о связанных с данной персоной организациях.

Для объекта «организация», кроме названий и контактных данных, приводятся тип организации, её правовой статус, даты создания и ликвидации, регион, политический сектор, сведения о лидере и о связанных органах власти и других организациях.

Объекты «документы», кроме библиографической информации, содержат связи с персонами, организациями, другими документами и указание на источник получения документа.

Полные тексты документов хранятся в файлах Системы и представляются пользователю в формате хранения.

Полные тексты документов поступают в Систему за счет импорта из действующей системы (<http://www.partinform.ru/pa98/>), из Интернета, из ресурсов ИНИОНа, а также за счет самостоятельной подготовки (оцифровка и последующая обработка).

Все атрибуты по требованиям к их значениям делятся на три категории:

- неконтролируемые – отсутствует формальный контроль их значений, например, заглавие документа;
- контролируемые по формату, например, даты;
- контролируемые по словарю – значение должно быть только из соответствующего словаря.

Для атрибутов, контролируемых по формату, должны быть разработаны требования к значениям и средства контроля, для атрибутов, контролируемых по словарям – словари значений, средства их ведения и средства контроля.

В ИС «Партархив» используются словари по следующим направлениям:

- политический сектор (персоны, организации),
- регион (персоны, организации),
- тип организации,
- правовой статус (организации),
- вид документа,
- темы, сгруппированные по предметным областям.

Представление объектов конечному пользователю. Объект «Персона» представляет собой веб-страницу, содержащую заполненные значения атрибутов данного типа. В качестве заголовка страницы используется имя персоны. Рядом с заголовком размещается фотография персоны.

Объект «Организация» представляется веб-страницей, содержащей заполненные значения атрибутов данного типа. Состав отображаемых полей зависит от типа организации. В качестве заголовка страницы используется название организации и её тип. Рядом с заголовком размещается логотип организации.

Если текстовый документ заимствован с внешнего сайта, то он представляется пользователю веб-страницей (html) с разметкой оригинала. При этом инкорпорированные в оригинал текст мультимедийные объекты передаются в тексте документа со своими заглавиями и ссылкой на соответствующий файл. Ссылка идет от заглавия и при клике на него производится переход на соответствующий сайт. Если сайта уже нет в сети, то вызывается соответствующий мультимедийный объект из ИС.

Если текстовый документ получен в результате самостоятельной обработки, то он представляется пользователю в двухслойном формате (изображение + текст; jpg + html). Формат (слоя) представления выбирает пользователь. При этом слой изображения полностью воспроизводит исходный текст, включая фотографии, диаграммы, графики и т.п., которые в текстовом слое (html) не воспроизводятся.

Если текстовый документ заимствован из ресурсов ИНИОНа в электронном виде, то он обрабатывается так же, как и заимствованный из Интернета, а при необходимости – преобразуется в формат html и формат изображения (jpg) сохраняется.

Традиционные документы ИНИОН обрабатываются так же, как и документы самостоятельной обработки.

Общая архитектура

Сайт Системы состоит из пяти разделов: Персоны, Организации, Регионы, Политический сектор, Документы.

Раздел выбирается на любой странице сайта, каждая главная страница содержит его описание.

Каждая страница раздела, включая главную, содержит в правой части навигационную панель с соответствующими указателями.

При активации того или иного указателя в левой части страницы открывается его содержание.

Если выбирается конкретный объект, то в указателе соответствующая статья открывается в новой вкладке.

Указатели

Основное средство навигации в Системе, представляющее собой различные упорядоченные подмножества объектов, сформированные по определенным признакам, – это Указатели.

Каждый объект в указателе имеет своё имя и если есть – его синонимы, например:

«КПРФ» — «Коммунистическая партия Российской Федерации»;

«Лавлинский Евгений Николаевич» — «Прилепин Захар».

В качестве синонимов рассматриваются псевдонимы, аббревиатуры организаций и ранее действовавшие названия организаций, вторые заглавия документов и т.д.

Омонимия в указателе устраняется путем введения соответствующих цифровых индексов.

Пользователю предоставляются следующие указатели:

- Персон — общий;
- Персон — по политическому сектору;
- Персон — по региону;
- Организаций — общий;
- Организаций — по типу;
- Организаций — по политическому сектору;
- Организаций — по региону;
- Организаций — по годам создания и ликвидации;
- Текстовых документов — по заглавиям.

Действующие организации в указателях должны быть соответствующим образом помечены. Статьи, вызываемые из указателей, открываются в новой вкладке; имена объектов располагаются в алфавитном порядке в две-три колонки, что обеспечивает быстрый переход к именам, начинающимся с выбранной буквы.

Функциональность

ИС «Партархив» включают пользовательские и административные функции. К пользовательским функциям относятся навигация, поиск, экспорт информации, обратная связь, а также разные виды справок и помощи. Административные функции, выполняемые персоналом системы, включают отбор информации, верификацию, обработку и ввод данных.

Навигация осуществляется через структуру сайта и указатели.

Поиск может быть атрибутный, лексический в текстах документов, а также комбинированный. Лексический поиск производится с учетом морфологии русского языка. Результаты поиска могут быть отсортированы по авторам, заглавиям, датам публикации документов. Экспорт информации предусматривает возможность создать pdf статьи для объектов «персона» и «организация», а также скачивать документы.

На сайте Системы должна быть предусмотрена обратная связь с пользователем, а также предоставление ему справочной информации (описания, помощь, подсказки).

Технология обработки и ввода информации

Подготовка информации для объектов «персона» и «организация», а также для полных текстов, осуществляется в специальном редакторе. В этом же редакторе вводятся атрибуты объектов, на основе которых должна выполняться проверка объектов на возможное дублирование – каждый новый объект сравнивается по определенным критериям с уже существующими. При наличии потенциально-го дублирования администратор Системы принимает решение о дальнейшей обработке этого объекта или отказе от неё.

Подготовка полных текстов при самостоятельной обработке исходного (печатного) текста включает следующие основные операции:

- сканирование (получение изображения всех страниц издания);
- формирование графического образа издания, включающее устранение «грязи» в оригинал, отбеливание, восстановление полиграфических погрешностей, выравнивание, приведение к установленным размерам, редактирование цветности и графики иллюстративного материала и др.;
- распознавание (получение из графического образа издания алфавитно-цифрового представления);
- формирование текстового образа издания, включая спеллчекинг;
- разметка структуры издания и входящих в него произведений, создание файла описи издания;
- создание графического индекса (файла соответствия графического и текстового образов издания);
- формирование загрузочного пакета информации в электронную библиотеку (включая индексирование и пополнение навигационных средств).

После завершения обработки информация загружается для окончательной проверки в тестовую систему, являющуюся точной копией публичной. Тестовая система доступна только специалистам, подготавливающим информацию.

Если результат проверки отрицательный, то информация возвращается на доработку. После положительного результата проверки информация может быть загружена в публичную систему.

Общие требования к эксплуатации

Администратор системы имеет право удалять любые объекты, при этом безусловно должна соблюдаться целостность баз данных.

При загрузке нового объекта все указатели обновляются автоматически, происходит переиндексирование для поиска.

В процессе эксплуатации выполняется регулярная проверка Системы и, в случае необходимости, редактирование гипертекстовых ссылок на внешние информационные ресурсы.

Вся информация и, прежде всего, полные тексты документов подлежат архивному хранению.

Для внесения изменений в структуру сайта, размещения сопроводительной информации используется контент-менеджер системы (CMS).

На первом этапе работы Системы выполняются:

- разработка базового программного и технологического обеспечения, создание сервера;
- разработка дизайна;
- подготовка в ограниченном объеме качественных данных для загрузки;
- загрузка подготовленных данных на публичный сервер.

Основной результат первого этапа: бета-версия Системы, находящаяся в публичном сетевом доступе.

Срок исполнения первого этапа определен как июль 2024 г.

На втором этапе происходит:

- завершение разработки программного и технологического обеспечения;
- тестирование бета-версии системы и проведение необходимых доработок;
- штатная подготовка данных для загрузки в систему;
- штатная загрузка подготовленных данных в систему;
- сдача системы в штатную эксплуатацию.

Основной результат второго этапа: штатная система, находящаяся в публичном сетевом доступе.

ЗАКЛЮЧЕНИЕ: ПЕРСПЕКТИВЫ РАЗВИТИЯ ИС «ПАРТАРХИВ»

Определенные в настоящее время параметры информационной системы «Партархив» могут быть скорректированы с учетом реальной разработки. Так, сегодня еще не определен характер информационных ресурсов, которые будут загружены в Систему из ресурсов ИИОН РАН – как электронных, так и традиционных. Ожидается определенные трудности при миграции больших массивов данных из действующей версии. Очевидно, что потребуется разработка или

корректировка словарей системы, о которых в настоящей статье почти ничего не сказано. Мы надеемся на расширение функциональности описываемой Системы, которая должна обеспечивать не только информационные и архивные функции, но и служить информационной базой для проведения научных политологических исследований.

СПИСОК ЛИТЕРАТУРЫ

1. Коргунюк Ю.Г. Структура электоральных размежеваний по итогам думских выборов 2021 года: тенденции трансформации и подходы к изучению // Электоральная политика. – 2023. – № 1. – URL: <http://electoralpolitics.org/ru/articles/struktura-elektoralnykh-razmezhevaniy-po-itogam-dumskikh-vyborov-2021-goda/>
2. Алексеев Д.В., Коргунюк Ю.Г., Помигуев И.А. Размежевания в российском законодательном процессе: опыт факторного и дискурс-анализа // Дискурс-Пи. – 2022. – Т. 19, № 3. – С. 10–35. DOI: https://doi.org/10.17506/18179568_2022_19_3_10/
3. Коргунюк Ю.Г. Советское прошлое и межпартийное противостояние : опыт использования количественных методов // Политическая наука. – 2021. – № 1. – С. 274–300. – DOI: <http://www.doi.org/10.31249/poln/2021.01.12/>
4. Akimova M., Belousova A., Pilshchikov I., Polilova V. CPCL: A Multilingual Parallel Corpus of Poetic Texts and New Perspectives for Comparative Literary Studies // Parallel Corpora as Digital Resources and Their Applications: [Part of DHN2020 (Digital Humanities in the Nordic Countries: 5th Conference). Riga: University of Latvia].
5. Пильщикова И.А. Семь бесед о филологии и Digital Humanities: Интервью и дискуссии (2015–2011) / общ ред и сост. В.С. Полиловой. – Москва: Изд-во Московского университета, 2022. – 190 с.

Материал поступил в редакцию 29.09.23.

Сведения об авторах

АНТОПОЛЬСКИЙ Александр Борисович – доктор технических наук, главный научный сотрудник ИИОН РАН, Москва.
e-mail: ale5695@yandex.ru

ВИГУРСКИЙ Константин Владимирович – кандидат технических наук, главный специалист ИИОН РАН
e-mail: vigur06@mail.ru

КОРГУНЮК Юрий Григорьевич – доктор политических наук, заведующий отделом ИИОН РАН
e-mail: partinform@mail.ru

УВАЖАЕМЫЕ КОЛЛЕГИ!

ВИНТИ РАН предлагает Вашему вниманию База данных (БД) ВИНТИ РАН в режиме online

База данных (БД) ВИНТИ РАН — Федеральная база отечественных и зарубежных публикаций по естественным, точным и техническим наукам. Генерируется с 1981 г., обновляется ежемесячно, пополнение составляет более 600 000 документов в год.

БД ВИНТИ РАН включает 26 тематических фрагментов, состоящих из более чем 190 разделов.

Документы БД содержат библиографию, ключевые слова, рубрики и реферат первоисточника.

На основе БД ВИНТИ пользователям доступны следующие продукты:

- online доступ к базе данных круглосуточно, без выходных;
- выполнение тематического поиска специалистом ВИНТИ по запросу заказчика;
- по заявкам предоставляются любые наборы тематических фрагментов БД ВИНТИ или их разделов на любых видах электронных носителей, или через FTP-сервер;
- для ознакомления с возможностями поиска имеется демо-версия базы данных bd.viniti.ru.

База данных ВИНТИ зарегистрирована Российским агентством по правовой охране программ для ЭВМ, баз данных и топологий интегральных микросхем (РосАПО) (Свидетельство № 960034 от 23.09.1996г.)

Подробную информацию Вы можете получить:

Адрес: 125190, Россия, Москва, ул. Усиевича, 20, ВИНТИ РАН

Телефон: 8 499-152-54-81

E-mail: feo@viniti.ru

УВАЖАЕМЫЕ КОЛЛЕГИ!

ВИНИТИ РАН предлагает Вашему вниманию Реферативный Журнал в электронной форме

РЖ в электронной форме (ЭлРЖ) выпускается по всем разделам естественных, технических и точных наук.

Каждый номер ЭлРЖ является полным аналогом печатного номера РЖ по составу описаний документов, их оформлению и расположению. Он сопровождается оглавлением, указателями.

ЭлРЖ представляет собой информационную систему, снабженную поисковым аппаратом и позволяющую пользователю на персональном компьютере:

- читать номер РЖ, последовательно листая рефераты;
- просматривать рефераты отдельных разделов по оглавлению;
- обращаться к рефератам по указателям авторов, источников, ключевых слов;
- проводить поиск документов по словам и словосочетаниям;
- выводить текст описаний документов во внешний файл.

ЭлРЖ могут быть:

- записаны на DVD-ROM;
- передаваться через FTP-сервер (клиенту предоставляется логин и пароль с доступом к FTP-серверу ВИНИТИ, с которого он скачивает заказанные журналы).

Электронные реферативные журналы можно заказать за текущий год с любого номера, а также за предыдущие годы.

Подробную информацию Вы можете получить:

Адрес: 125190, Россия, Москва, ул. Усиевича, 20, ВИНИТИ РАН

Телефон: 8 499-152-62-11

E-mail: feo@viniti.ru

ВИНИТИ РАН

Центр научно-информационного обслуживания

Информационные услуги, предоставляемые ЦНИО ВИНИТИ РАН:

- проведение тематического поиска и консультации поисковых экспертов;
- подготовка списков научной литературы;
- подбор, копирование полнотекстовых материалов из первоисточников на бумажном носителе и в электронном виде;
- библиометрическая оценка публикационной активности исследователей и научных организаций с использованием российских и зарубежных баз данных;
- информационное обеспечение информационно-аналитической деятельности по подготовке и предоставлению аналитических обзоров и других научных материалов.

ВИНИТИ РАН располагает следующими информационными ресурсами:

- фондом НТЛ, включающим более 2,5 млн. отечественных и иностранных журналов, книг, депонированных рукописей, авторефератов диссертаций и другой научной литературы, ретроспектива – с 1991 года;
- базами данных и Интернет-ресурсами: БД ВИНИТИ (разработка ВИНИТИ), БД SCOPUS, БД Questel (патенты) и другими реферативными ресурсами;
- полнотекстовыми электронными ресурсами (статьи, патенты, материалы конференций).

Ознакомиться с информацией о доступных полнотекстовых и реферативных ресурсах можно на сайте ВИНИТИ РАН www.viniti.ru

К услугам пользователей – **Электронный Каталог ВИНИТИ** <http://catalog.viniti.ru> и **служба электронной доставки документов.**

Осуществляется платное информационное обслуживание по разовым заказам и на договорной основе с предоставлением всех необходимых финансовых документов.

Проводится индивидуальное обслуживание пользователей в читальном зале ЦНИО ВИНИТИ РАН.

Подробную информацию Вы можете получить:

Адрес: 125190, Россия, г. Москва, ул. Усиевича, 20, ВИНИТИ РАН;

Телефоны: 499-155-42-17, 499-155-42-43;

E-mail: cnio@viniti.ru