

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 1. ОРГАНИЗАЦИЯ И МЕТОДИКА
ИНФОРМАЦИОННОЙ РАБОТЫ

ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 10

Москва 2015

ОБЩИЙ РАЗДЕЛ

УДК 004.738 : 316.658 : 316.776.23

О. В. Сютюренко

Сетевые технологии информационного противоборства и манипуляции общественным сознанием

Рассматривается влияние новых информационных технологий на развитие социокоммуникаций и социосистем. Анализируются свойства и характеристики современной информационной среды, исследуются факторы и технологии информационного противоборства и конфронтации. Показана роль информационно-коммуникационных технологий в функционировании сетевых социальных общественных движений. Раскрываются средства и методы использования СМИ, социальных сетей и деструктивных сетевых технологий для манипулирования групповым поведением и трансформации общественного сознания. Рассматривается потенциал социодинамики как нового инструмента целевого воздействия и влияния на социосистемы. Оцениваются потенциальные угрозы и риски развития систем и сетей телекоммуникаций. Сформулированы некоторые выводы и актуальные задачи по предотвращению нарастания информационной конфронтации и противодействию скрытым воздействиям и угрозам в отечественной социально-экономической и научно-технической сферах.

Ключевые слова: информационная среда, информационные технологии, медиавирусы социокоммуникации, суперкомпьютинг, Большие данные, информационная

ВВЕДЕНИЕ

Информация становится одним из основных экономических ресурсов и структурирующим социальным фактором, под воздействием которого меняются формы экономической деятельности, виды и типы промышленного производства, социальные взаимоотношения.

Информационные методы и технологии все глубже проникают практически во все сферы человеческой деятельности. Информатизация, конвергенция компьютерных, телекоммуникационных технологий и мультимедиа обеспечивают принципиально новый уровень цивилизационного развития. Вместе с тем информатизация общества порождает много новых и тревожных проблем в социальной сфере. Возрастающее использование информационных технологий существенно влияет на экономические и социальные процессы, изменяет отношения между властью и гражданами, порождает новые необычные проблемы (политические манипуляции через Интернет) и знаменует наступление эпохи «компьютеризованной политики, экономики, культуры». Быстрое развитие информационно-коммуникационных технологий, расширение сетевой информационной среды, изначальная интерактивность Интернета и появление так называемых социальных сетей влечет появление дополнительных рисков, угроз общественной стабильности, устойчивости и безопасности на национальном и международном уровнях.

Глобальное сетевое пространство становится ареной информационного противоборства. Основной (и конечной) целью информационного противоборства является разработка и практическое использование синтетических технологий управления групповым и массовым сознанием в странах-конкурентах (странах-противниках). Информация способна не только отражать реальность, она может активировать создание новой реальности, которой до этого не было. Модель действительности может подталкивать саму действительность в ином направлении, чем это есть сегодня, например, в финансовой сфере, как это показано в книге Д. Сороса «Кризис капитализма». Война информационная перерастает в войну смысловую. Возрастает роль социальных наук. Развитие Интернета и сетевых технологий, а также континуум новых технологий (когнитивные вычисления, когнитивная психология, социодинамика, бихевиоризм), обуславливают новую роль общества в геополитике. Общество изнутри действует на поведение государства во внешней политике, формируя поведение элит «снизу». Поведение общества внутри государства становится инструментарным фактором, на который можно воздействовать извне, тем самым формируя поведение государства на внешнеполитической арене. Главенствующую роль играют сети, созданные на основе таких исключительно социологических явлений, как социальное действие и социальное взаимодействие.

НОВЫЕ ФЕНОМЕНЫ НОВОЙ ИНФОРМАЦИОННОЙ СРЕДЫ

Анализ сетевой, цифровой среды, медиа- и социокommunikаций позволяет выделить основные черты, свойства (феномены) современного глобального информационного пространства:

- лавинообразный рост мирового объема информации;
- быстрый рост глобальной телекоммуникационной сети. По прогнозам компании Cisco число IP-адресов превысит 50 млрд к 2020 г.;
- трансграничность Интернета, высокий уровень анонимности и потенциальная бесследность при реализации угроз и несанкционированных действий, сложность контроля сетевого пространства и актуализация проблем информационного суверенитета;
- устойчивые тенденции прогрессирующего роста компьютерной преступности, распространения пропагандистских материалов преступных организаций;
- возникновение феномена консциентальных и сетевых войн, нарастание информационного противоборства с использованием социальных сетей и медиа;
- расширение социокommunikаций и рост сетевых общественных движений, возрастание роли социальных медиа в организации «цветных» революций;
- появление и распространение альтернативных, так называемых цифровых валют типа Биткойн (Bitcoin);
- смена в социосфере дефицита информации дефицитом внимания в условиях развития СМИ и социальных медиа, поскольку информации стало слишком много, роль факта уменьшается, возрастает роль интерпретации. Как следствие – возникновение феномена трансмедиа;
- отсутствие эффективного международного правового регулирования и актуализация проблем цифрового неравенства.

Постиндустриальные технологии делают акцент на развитии коммуникаций, передаче информации, и здесь ключевым элементом и средой распространения этой информации является сеть (в ее реальном и виртуальном аспектах). Сеть представляет собой полицентричное неиерархизированное информационное пространство. Особенностью информационного пространства XXI в. является возможность производства, хранения, передачи и обработки информации в любом количестве и с огромной скоростью. Также особенностью постиндустриального общества является то, что количество информации достигло такого объема, что ни один нормальный человек не способен не то чтобы усваивать этот объем, но даже хоть как-то в нем ориентироваться. По оценкам International Data Corporation (IDC), мировой объем информации удваивается каждые два года. По данным компании Cisco, только мобильный трафик дан-

ных в мире в начале 2013 г. достиг ~ 900 петабайт (один петабайт = 10 в 15 степени байт) в месяц (к 2016 г. доля пользователей, применяющих гаджеты для выхода в Интернет, составит ~ 48%). К 2016 г. ежегодный объем глобального IP-трафика составит 1,3 зеттабайта (один зеттабайт = миллиарду гигабайт). Кроме того, зачастую умышленно создаются условия для того, чтобы поток информации постоянно усиливался, а способность усвоить и критически осмыслить полученную информацию каждым индивидуумом соответственно снижалась. Главное, что дефицит информации сменился дефицитом внимания, поскольку информации стало намного больше. И теперь приходится управлять именно вниманием массового сознания, удерживая его на нужных объектах и ситуациях как можно дольше. Массовое сознание достаточно инерционно, поэтому кампании по его изменению как правило требуют большого времени и ресурсов. Особенно, когда это связано хотя бы с минимальным противодействием.

Современный Интернет имеет техническую и гуманитарную составляющую. По технической линии развитые страны получают сотни атак ежедневно, однако, в основном - это хакерство, технический или экономический шпионаж. В 2014 г. было пресечено 75 млн попыток атак на официальные сайты органов власти в России. Помимо этого было выявлено свыше 25 тыс. интернет-ресурсов с публикациями, нарушающими закон, а более полутора тысяч сайтов, относящихся к экстремистским, прекратили свою работу. Интернет в гуманитарном аспекте – это социальные медиа, которые покрывают собой большой сегмент информационного (в основном нецензурируемого) пространства. Роль медиа в современном мире постоянно возрастает. Медиа стали инструментом государственного управления со стороны власти и давления на власть со стороны оппозиции. Интернет облегчает социальные протесты, делая анонимным выражение мнений и координацию коллективных действий, что может вести к эффекту домино, как это видно на примере так называемых «бархатных» и «цветных» революций.

Сетевое противоборство осуществляется на информационном уровне, уровне смыслов и идей, с использованием информационных технологий, неправительственных организаций, блогерского сообщества, с подключением журналистов, политиков, СМИ. Это перманентный, многоуровневый процесс, направленный на формирование модели поведения (группового, массового) и целевой корректировки системы принятия решений. И внешние, и внутренние силы пытаются влиять на системы принятия решений. В более конкретном выражении это принимает вид влияния на индивидуальное или массовое сознание. Воздействия по смене и трансформации типов имидж-идентификаций и аутентизации весьма эффективно осуществляют средства массовой информации, в первую очередь телевидение – самый влиятельный тип СМИ в России, формирующий взгляды аудитории. Телевидение как инструмент управления и власти выполняет задачу интерпретации действительности (чем ярче медиакартинка, тем более верным и правильным запоминается событие), снятия напряженности, отвлече-

чения и развлечения. В противном случае в головах большинства людей воцарился бы хаос. Во время промышленной революции в Англии стремительно возросло потребление джина (и шоколада в США во времена Великой депрессии), который выступил технологией для снятия напряжения. Сегодняшним вариантом этой технологии стало потребление телепродукции. Здесь следует отметить уже упоминавшийся феномен трансмедиа, когда осуществляется распределение сюжета-темы по разным платформам, чтобы получить в результате мультиплатформенное отражение-рассказывание. Это важно, так как одни люди сидят в одних информационно-новостных нишах, другие – в других. Все возможные средства и Интернет, и масс-медиа используются сразу, как в политических, так и в коммерческих рекламных целях. Все шире распространяются и видеоигры, которые также несут в себе потенциал изменений, внедряющихся в массовое сознание. Сетевое противоборство (сетевая война) – это эффективное средство воздействия, за которым следуют социальные трансформации и реальные геополитические изменения. Сетевые войны – это явление, которое не всегда связано с Интернетом, но всегда связано с таким понятием, как сеть. В сети участники – субъекты, объекты, акторы, узлы – выстраиваются не вертикально, как в традиционных обществах, а горизонтально (при этом лидерство распределяется на большое число активистов-акторов). Сеть нелинейна и многосложна, все ее узлы связаны между собой бесконечным количеством связей. Современные социальные сети облегчают формирование различных, в том числе и протестных, движений (в том числе «бархатных» и «цветных» революций), ускоряют политические изменения; делают слабые связи сильнее; влияют на организацию целеориентированных новостных циклов в СМИ. Дискурсы и дискуссии в Интернете облегчают, закрепляют и усиливают радикальные суждения и взгляды. Для достижения эффекта влияния важно добиться не только простого реагирования на изменения, но необходимы целенаправленные действия по активному формированию (корректировке) транспарентной социальной сетевой среды. В последнее десятилетие следует констатировать выраженный тренд роста: а) деструктивных сетевых социальных структур; б) искусственно создаваемых сетевых ресурсов, используемых для организации информационных атак; в) подложных (клонированных) ресурсов сети Интернет, которые могут быть использованы для дезинформации широкого круга пользователей.

МЕДИАВИРУСЫ «ЦВЕТНЫХ» РЕВОЛЮЦИЙ

В рамках деструктивных сетевых технологий все более широко используются медиавирусы (информационно-психические вирусы), так называемые мемы, ставшие основным фактором воздействия в относительно недавних событиях «арабской весны» и на Украине. Мем можно рассматривать как своего рода единицу подражания. Примерами мемов являются лозунги, определенный тип одежды (мода в целом), конкретные способы сделать что-то, молодежный

сленг и воззрения, виды развлечений и т.д. Медиа-вирусы подробно анализируются в ряде отечественных и зарубежных публикаций, где мемы рассматриваются в связи с их способностью распространяться по Интернету и каналам СМИ, вызывая социально значимые процессы и последствия. Мем – это максимально доступный, деструктивный симулякр. Часто мем (группа мемов) формирует то, что определяется как *настроение в средах*. Главное условие живучести мема – быстрое, массовое и повсеместное распространение. Мемы по сути являются темпоральными объектами, которые существуют только в процессе своей реализации. Эра темпоральных объектов наступила с появлением кинематографа, радио и телевидения. Уровень темпоральности зависит от того, насколько все чувства индивидуума вовлечены в процесс восприятия предлагаемого сюжета (внешнего дискурса, проблематики, привносимых ценностей и т.п.). Максимальный уровень темпоральности достигается в процессе интерактивного взаимодействия (сетевые компьютерные игры, социальные сети, интерактивное ТВ). Уровень темпоральности влияет на самоидентификацию и интенсивность трансформации сознания. Виртуально человек может чувствовать себя более значимым и свободным, чем в реальной жизни. Иллюзорное чувство виртуальной свободы в соединении с максимальной вовлеченностью в темпоральный объект обуславливает возникновение эффекта группового управления поведением в сети. Этот феномен инициирует и полностью определяет процессы «цветных» революций. Под влиянием социальных медиа, прямых телевизионных трансляций, массы зрителей (пользователей Интернета) принимают как свое время «цветных» революций. Таким образом, возможно создание в сети виртуального дубля гражданского общества, чтобы противопоставить его власти, которая в эту сеть не входит. Поскольку сеть является дублем, подделкой гражданского общества, то и все действия ее симулятивны и, в основном, деструктивны. Для того чтобы квазитолпа превратилась в субъект действия, одних процессов самоорганизации недостаточно. Должно осуществляться внешнее управление, тайминг, логистика. Создание мемов – это целая отрасль, которая должна учитывать множество нюансов (культурные особенности сред, механизмы работы медиа, форматы СМИ в разных странах и т.д.). Достоверность факта здесь не является столь важным параметром, как, например, в научных социальных сетях. Факт вторичен, первичным является требуемый образ-символ и интерпретация факта.

Социальная активность и массовая ретрансляция мемов обычно инициируются из ограниченного количества центров: 2% пользователей Twitter создают 60% всего контента, распространенного в сети, а 5% пользователей создают уже 75% контента. Это – специально ангажированные, оплачиваемые «пользователи» сети. Потом, когда большое количество реальных пользователей (принимающих это как *свое*) подключается к процессу, ретрансляция продолжается самопроизвольно. Если мем оказывается живучим, то дальше это явление самоиндуцируется в своем распространении. При этом следует отметить, что

сеть состоит как из самих пользователей, так и из виртуалов и так называемых ботов – виртуальных юзеров, находящихся в управлении сетевыми роботами, подконтрольными ангажированным блогерам. Таким образом формируется *настроение в средах* и, в итоге, то или иное общественное мнение или идеологические установки (технологии *sock puppet revolution* и *twitter-революций*). В 2015 г. на базе чешского филиала Radio Free Europe/Radio Liberty (Радио «Свобода», финансируется федеральным бюджетом США) будет создан цифровой медиадепартамент, DIGIM. В этом подразделении будут работать специалисты по социальным сетям, их задача – «противостоять дезинформации в российской медиасфере посредством различных соцмедиаплатформ (в частности, Facebook, Twitter, «ВКонтакте» и «Одноклассники»)). По оценкам отечественных и зарубежных специалистов, сеть Интернет как среда поиска и обмена информацией все более приобретает роль глобального инструмента ведения психологической войны. В целом все деструктивные сетевые технологии являются частью и развитием так называемой «стратегии не прямых действий», детально представленной в книге Базила Л. Гарта с одноименным названием.

Безусловно информационно-коммуникационные технологии повышают уровень функционирования социальных движений (сетевых общественных движений). Они сделали эти явления эффективнее, поскольку качественно повысили их масштабность и мобилизационные возможности. Современное информационное общество представляет собой идеальную среду для создания, функционирования и использования сетей. При этом следует подчеркнуть, что социальные сети, как следует из названия, это явление социальное, оно предусматривает искусственно сконструированную коллективную субъектность, представляющую часть общества, и отражает его среды, социальные группы, классы. Следовательно, сетевые общественные движения и социальные сети неоднородны, а их участники могут переконfigurироваться в зависимости от социального статуса, что делает такие сети довольно гибкими. Таким образом, социокommunikации (и социальные сети в частности) дают возможность поддерживать имеющееся состояние социосистемы; переводить социосистему в новое состояние; активировать или блокировать определенные параметры социосистемы.

СОЦИОДИНАМИКА КАК ИНСТРУМЕНТ УПРАВЛЕНИЯ МАССОВЫМ СОЗНАНИЕМ

В последние годы теория и практика социодинамики дополнились новым научным направлением – социодинамикой, которое обобщает эмпирические закономерности, полученные в результате: а) применения технологий Big Data (Большие данные) к огромным массивам информации, содержащейся в банках данных крупнейших социальных платформ, таких как Twitter, Facebook, Google и др. (60% этой информации носит неструктурированный, в основном текстовый характер); б) использования методов многомерного математического и статистического анализа, когнитивных и поведенческих наук. Эти эм-

пирические закономерности используются для обработки практических методов внешнего воздействия, управления и манипулирования различными сообществами и социальными группами. Появились эффективные разработки, ориентированные на неявный сбор данных и скрытое управление групповым поведением коллективов большой размерности. Суперкомпьютинг (вычислительная основа технологии Big Data) позволяет автоматически сегментировать и обрабатывать в режиме реального времени огромные объемы информации. Суммируя полученный контент и выделяя в нем ключевые смысловые группы, можно получить информацию в различных разрезах об интересах и настроениях всех субъектов исследуемого сегмента. Анализ всего объема данных позволяет определить, какие настроения и ценностные ориентации преобладают в исследуемом социальном сегменте и, следовательно, каким образом информационно можно проникнуть в это сообщество (в том числе с использованием СМИ), с какими смысловыми и мировоззренческими посылами можно формировать нужный контекст в том или ином обществе или государстве. Фактически социодинамика - это новый информационный инструментарий управления чужими смыслами, входа на «чужую» территорию. Поведение общества (или крупных социальных групп) внутри государства становится инструментальным фактором, на который можно воздействовать извне, тем самым формируя поведение государства на внутриполитической и внешнеполитической арене. Таким образом, развитие методологии и методов социодинамики, использование технологий Big Data в перспективе позволит отслеживать динамику социальных взаимодействий (и их происхождение), эффективно управлять любыми социальными сетями, сообществами и целевыми группами при условии знания закономерностей их формирования и динамики, а также характеристик групп любых масштабов и структур, что, безусловно, породит новые риски возникновения конфликтов и нестабильности для социума. Следует также отметить, что технологии Big Data обеспечили создание новых, на порядок более эффективных, чем раньше, методов прогнозирования экономических, научно-технических, политических и социальных процессов. По отношению к социосистемам и социокоммуникациям эти методы – принципиально новый инструмент информационного противоборства.

РИСКИ РАЗВИТИЯ СИСТЕМ И СЕТЕЙ ТЕЛЕКОММУНИКАЦИЙ

Развитие информационно-технологической среды мировой экономики сопровождается нарастанием угроз функционированию сложных информационных систем управления промышленной, энергетической, информационной инфраструктурой. Устойчивая тенденция расширения сетевого пространства, увеличения числа пользователей Интернета (по различным оценкам до ~70% населения всех стран в 2015 г.) также влечет за собой рост проблемных рисков для большинства государств. Уязвимыми элементами сети Интернет являются система присвоения доменных

имен и система передачи интернет-трафика. Присвоение доменных имен осуществляет неправительственная организация ICANN (Ассоциация по присвоению имен и портов Интернета, США). Там же принимаются решения и по регистрации доменов верхнего уровня, таких как .net, .com, .info, .org, а также национальных доменов, куда входит и доменная зона .ru. Гипотетически из этого следует, что как ICANN зарегистрировала национальную доменную зону, так она может ее и закрыть. В этом случае все сайты и вся пересылка почты для данной доменной зоны перестанут быть доступны (на неопределенное время). Существует аналогичная уязвимость и интернет-трафика. Передача интернет-трафика со стороны тех или иных национальных зон может быть прекращена в любой критический момент. Так как американские провайдеры к интернет-трафику имеют полный доступ, они могут предоставлять его своему правительству, как и крупнейшие поставщики интернет-услуг, в том числе услуг электронной почты типа Gmail, Postfix или Hotmail.

Следует констатировать недостаточную устойчивость работы Интернета на территории России в случае недружественных целенаправленных действий. Ключевые элементы Интернета находятся за пределами РФ. Доступ (к сайту) может быть заблокирован на любом этапе: корневые серверы могут неправильно перевести адрес сайта в машинную форму, не отвечать на запросы, физически отключить сегмент сети от Интернета. Таким образом RuNet может быть заблокирован целиком. Прецедент – отключение Сирии от Интернета в 2012 г. Именно в силу этих причин объясняется напряженность, возникающая из-за потенциальных угроз отключения российского сегмента Интернета от международных платежных систем Visa и Mastercard, а также от международной межбанковской системы передачи информации и совершения платежей SWIFT. Некоторые страны, например Китай, осознавая риски уязвимости интернет-трафика, приступили к созданию национальных сегментов Интернета. Китай предпринимает и другие меры, чтобы обезопасить себя от деструктивного вторжения и неконтролируемого доступа к информационным ресурсам. Ставя вопрос о полном отказе от пользования системами Microsoft, он добился передачи ему исходного программного кода операционной системы Windows, а также исходных текстов программного обеспечения маршрутизаторов фирмы Cisco, которые обеспечивают работу большинства мировых сетей и серверов (и, кстати, производятся в Китае). Следует также отметить, что практически все ведущие компании, специализирующиеся на разработке программных решений по информационной безопасности сети Интернет, имеют американскую юрисдикцию так же как и все крупнейшие провайдеры (Twitter, Google, Amazon, eBay, Facebook и др.).

Постоянным и высоким фактором риска в сфере телекоммуникаций, энергетики, промышленности является широкое использование импортной электронной элементной базы. По некоторым оценкам существует потенциально высокий уровень рисков нарушения функционирования для ~90% отечествен-

ных энергосетей (невосстановимое отключение системы, перехват управления энергосистемой и т. п.) как из-за атак компьютерных вирусов (типа Dugu или Stuxnet), так и от внешних несанкционированных действий, осуществляемых за пределами их возможного обнаружения и идентификации.

В последние годы за рубежом активизировались разработки в области надежности и безопасности систем управления в различных областях промышленного и гражданского назначения. Наиболее важными документами в США в этой сфере являются «Национальная стратегия кибернетической безопасности» (The National Strategy to Secure Cyberspace) и «Национальная стратегия физической защиты критической инфраструктуры» (The National Strategy for The Physical Protection of Critical Infrastructures and Key Assets), в которых изложен подход к обеспечению взаимодействия всех участников программ по защите критической инфраструктуры государства. В феврале 2014 г. Белый дом объявил о создании Концепции кибербезопасности. Программа является добровольной, однако для ее функционирования создаются соответствующие подпрограммы для 16 секторов критической инфраструктуры под эгидой Министерства национальной безопасности (создано в 2003 г.), объединенные общей дорожной картой. Одной из основных задач Концепции является развитие частно-государственного партнерства в этой сфере.

ВЫВОДЫ. АКТУАЛЬНЫЕ ЗАДАЧИ

Сформулируем некоторые выводы, первоочередные и наиболее актуальные задачи по предотвращению нарастания информационной конфронтации, защиты национального информационного сетевого пространства и информационного суверенитета.

В мае 2015 г. в Москве Россия и Китай подписали соглашение по информационной безопасности и намерены совместно бороться с угрозами в киберпространстве и выступать с общих позиций на ключевых площадках, включая ООН и ITU (Международный союз электросвязи).

А. С высокой степенью вероятности можно прогнозировать, что такие факторы, как глобализация мировой экономики, внедрение новых информационных технологий (в том числе суперкомпьютинга и систем искусственного интеллекта), расширение мировой сети телекоммуникаций, развитие семантического Интернета и массмедиа, будут все более актуализировать проблему роста информационной конфронтации, информационного противоборства, нестабильности мирового сообщества.

В. Расширение информационного пространства, появление новых технологий, обеспечивающих возможность доминирования в различных сферах жизнедеятельности, совершенствование сетевых технологий воздействия на массовое сознание и скрытого управления групповым (и массовым) поведением, программирование деструктивных действий с использованием социальных сетей – все это на качественно новом уровне актуализирует проблемы информационной безопасности, информационного противоборства и информационного суве-

ренитета. Защита социума от новых информационных технологий входит сегодня в перечень наиболее актуальных задач.

С. Развитие глобальной телекоммуникационной инфраструктуры, когнитивные вычисления, экспертные системы, технологии Больших Данных позволяют моделировать экономические и социальные процессы и непосредственно воздействовать на них. Очевидна нарастающая актуальность многоаспектного исследования влияния новой глобальной информационной среды на межгосударственные отношения, развитие материальной и духовной жизни отдельного индивидуума и социума.

Д. С целью предотвращения нарастания информационной конфронтации (консциентальных, сетевых войн) необходимы согласованные усилия мирового сообщества, тесное сотрудничество и взаимодействие в сфере использования информационных технологий и информационной безопасности. Безотлагательной мерой является интернационализация управления Интернетом (например, под эгидой ООН), обеспечение на межгосударственном уровне цифрового равенства и информационного суверенитета стран – членов мирового сообщества.

Е. Для прогнозирования развития сложных социальных систем и противодействия скрытому (сетевому) управлению групповым поведением (коллективов различной размерности) необходима комплексная государственная Программа, а рамках которой будут осуществляться мониторинг и системный анализ информационного пространства, разработка методологии, математического аппарата и программных решений механизма противодействия различным угрозам и сетевой агрессии. Для реализации Программы должна быть разработана дорожная карта с взаимоувязкой сроков, комплекса высокотехнологичных мер, затрат ресурсов, а также принятия необходимых юридических актов на государственном уровне.

Ф. Для защиты национального сегмента сети (RuNet) в РФ необходимо создать: а) свою ключевую инфраструктуру Интернета, включая национальные корневые сервера, национальную систему маршрутно-адресной информации; б) свои электронные компоненты, оборудование, системное программное обеспечение (в качестве примера здесь можно отметить российский аналог СУБД Oracle – платформу «1С:Предприятие 8» компании «1С»).

ЗАКЛЮЧЕНИЕ

Следует констатировать, что проблематика, связанная с информатизацией общества, развитием информационных технологий, информационной безопасностью социума имеет междисциплинарный характер и больше не является прерогативой только лишь профессионалов и узких специалистов. Она становится наднациональной проблематикой всего мирового сообщества. С учетом нарастающих тенденций объединения информационных и вычислительных ресурсов многих стран в глобальные сети следует иметь в виду возможность трансформации традиционных проблем IT-угроз и рисков в проблему

компьютерного силового давления. Защита информационного пространства России является одним из приоритетных направлений обеспечения национальной безопасности. Поэтому стратегия развития информационной инфраструктуры и информационных технологий в нашей стране должна сочетать максимальное использование международного опыта, возможностей поиска, обмена, обработки информации в сетевых пространствах с минимизацией рисков негативного влияния извне на различные социальные группы и отечественные научно-технические и информационные ресурсы.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Смолян Г.Л. Сетевые информационные технологии и проблемы безопасности личности // Вестник РФФИ. – 1999. – № 3(17). – С. 63-68.
2. Рашкофф Д. Медиавирус. – М.: Ультракультура, 2003. – 363 с.
3. Siountiourenko O. The Problems of Providing Information Security: The Case of Information Infrastructure // Studies in Eastern Europe. Technological and Environmental Policy / ed. Gerhard Banse. – Berlin, 2007. – P. 163-178.
4. Стиглер Б. Индивидуация и массовое потребление. – URL: s357a.blogspot.com/2012/05/blog-post.html.
5. 50 фактов о Twitter, которые интересны всем. – URL: <http://twitt.ru/twitter-tips/50-facts.html>
6. Гарт Б.Л. Стратегия не прямых действий. – М.: Эксмо, Мидгард, 2008. – 462 с.
7. Дугин А.Г. Сетецентрические войны. Новая теория войны // Сетевые войны. Угрозы нового поколения. – М., 2009. – 200 с.
8. Дементьев И.О. Интеллектуализация глобальных телекоммуникационных систем с использованием семантических сетей // Искусственный интеллект: философия, методология, инновации: материалы 3-й Всероссийской конференции студентов, аспирантов и молодых ученых МИРЭА, 11-13 ноября 2009 г. – М., 2009. – С. 103-104.
9. Васенин В.А. Критическая энергетическая инфраструктура: кибертеррористическая угроза и средства противодействия // Информационные технологии. – 2009. – № 9. – С. 3-35.
10. Логинов Е.Л. Новые информационные технологии для контрольной деятельности в сфере государственного и корпоративного управления // Информационное общество. – 2011. – № 6. – С. 32-39.
11. Systems thinking and future studies (Systems thinking Ontario, 2013-03-21). – URL: coevolving.com/blogs/index.php/archive/systems-thinking-and-futures-studies-systems-thinking-ontario-2013-02-21/
12. Горохов В.Г., Сюнтюренько О.В. Технологические риски: информационные аспекты безопасности общества // Программные системы и вычислительные методы. – 2013. – № 4(5). – С. 344-353.
13. Сюнтюренько О.В. Макропроблемы информационной безопасности человека и общества // Научно-техническая информация. Сер. 1. – 2013. – № 6. – С. 11-14; Siountiourenko O. V. Macro-problems of the Information Security of Man and Society // Scientific and Technical Information Processing. – 2013 – Vol. 40, № 2. – P. 97-100.
14. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим. – М.: Изд-во Манн, Иванов и Фербер, 2014. – 240 с.
15. Коровин В.М. Третья мировая сетевая война. – СПб.: Питер, 2014. – 352 с.
16. Ларина Е.С., Овчинский В.С. Кибервойны XXI века. О чем умолчал Эдвард Сноуден. – М.: Книжный мир, 2014. – 351 с.
17. Выступление В.В. Путина на заседании коллегии ФСБ 26 марта 2015 г. – URL: <http://kremlin.ru/news/47963>.
18. Сюнтюренько О.В. Цифровая среда: тренды и риски развития // Научно-техническая информация. Сер. 1. – 2015. – № 2. – С.1-7; Syunturyenko O. V. The Digital Environment: The Trends and Risks of Development // Scientific and Technical Information Processing. – 2015 – Vol. 42, № 1. – P. 24-29.
19. США вводят информационные войска в российские социальные сети // Известия. – № 96, 14 апреля 2015 г. – URL: <https://news.mail.ru/politics/21711516/?frommail=1>.
20. Россия и Китай подписали соглашение по информационной безопасности // РИА Новости, 08 мая 2015 г. – URL: <http://ria.ru/defense.safety/20150508/1063359829.html/>

Материал поступил в редакцию 17.04.15.

Сведения об авторе

СЮНТЮРЕНКО Олег Васильевич – доктор технических наук, профессор, ведущий научный сотрудник ВНИТИ РАН, Москва
E-mail: olegasu@mail.ru

УДК 001.103.2 : 620.22 – 022.532

А.О. Еркимбаев, В.Ю. Зицерман, Г.А. Кобзев, М.С. Трахтенгерц

Универсальная система метаданных для характеристики наноматериалов^{*}

Содержится подробный анализ недавно разработанной системы идентификации наноматериалов и наноструктур произвольной природы. Система создана в результате совместной работы комиссии по численным данным в науке и технологии (CODATA) и Международного научного союза (ICSI). Показано, что по существу, разработанная система представляет собой расширенный набор метаданных, отражающих концептуальную схему научной дисциплины. Их особенности: широта в охвате всей совокупности свойств и признаков; возможность добавления новых рубрик, т. е. поддержка эволюции концептуальной схемы, что соответствует требованиям предметной области. Выявлены ограничения в возможностях представления определенных категорий наноматериалов, особенно макроскопических объектов. Показано, что их преодоление возможно путем активного привлечения существующих онтологий в области материаловедения.

Ключевые слова: метаданные, наноструктуры, наноматериалы, справочные данные

ВВЕДЕНИЕ

Метаданные (МД) или иначе «данные о данных» являются неперенным атрибутом любого хранилища информации: от обычных библиотек до компьютерных баз данных, Web-порталов и социальных сетей. Более точное определение, данное в докладе [1], трактует МД как «ключ к пониманию содержания информационной системы, обеспечивающий идентификацию, классификацию, определение местоположения, интеграцию и интерпретацию информации». В обычной практике библиотек или архивов МД описывают ресурс, предоставляя его адрес в хранилище или сети, сведения об авторах и издательстве, поисковые индексы и т.п. Роль и значимость МД резко возрастают при переходе к научным базам данных (БД), особенно в тех областях, где сбор, обработка и обмен данными занимают основное место в рабочем процессе, например, в био- и геоинформатике, исследованиях климата, астрономических наблюдениях и др. В этом случае структура МД, не ограничиваясь библиографическим описанием, включает множество атрибутов, характерных для соответствующей дисциплины, что позволяет пользователям БД вести узко ориентированный поиск, а создателям –

эффективный обмен данными с тематически близкими ресурсами. Существует обширный депозитарий (www.dcc.ac.uk/), где представлены стандарты МД для множества дисциплин, от астрономии до медицины и сельского хозяйства. Как правило, такие стандарты задают информационную модель предметной области, сочетающую общие сведения (название и участники исследования, его временные рамки, условия доступа, форматы и адреса наборов данных и т.п.) с МД, детализирующими предмет и метод исследования. К типовым атрибутам, раскрывающим научное содержание данных, относятся сведения об их происхождении (эксперимент, наблюдение, моделирование), качестве и неопределенности, единицах измерения, номенклатуре объектов, их свойствах и состояниях.

Многие из перечисленных понятий могут быть определены ссылкой на соответствующую позицию в классификаторах или контролируемых словарях, размещенных в сети. Так, в химии название вещества можно определить, ссылаясь на уникальный идентификатор – CAS Registry number, что исключает неоднозначность из-за синонимов и языковых различий. Ссылка на компендиум химической терминологии IUPAC [2], включенная в состав метаданных, позволяет точно идентифицировать почти все химические термины и концепции, относящиеся к структуре и

^{*} Работа выполнена при поддержке Российского научного фонда, грант № 14-50-00124.

свойствам вещества, механизмам реакций и т.п. Например, ссылка на URL <http://goldbook.iupac.org/R05396.html> однозначно определяет понятие *ring-opening polymerization*. Четкое выделение терминов ссылкой на соответствующий словарь, исключая произвол и неоднозначность в толковании, обеспечивает также использование программных агентов при поиске данных и интерпретации понятий [3].

В работе [4] авторы проанализировали богатые возможности МД в сопровождении определенной категории данных о свойствах веществ и материалов. Для теплофизических свойств веществ соответствующая система МД разработана в виде ThermoML, предметно ориентированной версии языка XML [5]. Предложенная многоуровневая иерархия метаданных позволила фактически сформулировать основу концептуальной модели, включающей всю номенклатуру объектов, их свойств, фазовых состояний, методов измерений и очистки образцов, типов неопределенности и др. Система способна описать данные для более чем 120 свойств из области теплофизики при табличной или аналитической формах представления и различном статусе данных (экспериментальные, расчетные, справочные). В качестве объектов могут рассматриваться чистые вещества, смеси постоянного состава, химические и фазовые реакции. Для представления данных разработана доступная в сети детальная схема [6]. Аналогичная схема языка XML предложена для представления и обмена данными в области материаловедения (язык MatML [7]), структура которых отягощена ссылками на условия производства, технологии и методы испытаний.

Естественно, что на повестке дня стоит задача распространить инструментарий МД на область нанотехнологий, учитывая беспрецедентный масштаб этой дисциплины и связанный с ним рост информационного потока в виде публикаций, БД, Web-порталов и др. Однако на пути систематизации этих данных возникает целый ряд проблем, обусловленных спецификой предмета и его местом в общей системе естественнонаучных и инженерных знаний. Анализ, выполненный в ряде исследований [8-10], выявил целую гамму особенностей, относящихся к наноматериалам (НМ):

- широта и нечеткие границы предметной области с интеграцией многих разделов физики, химии, биомедицины и др.
- быстрая эволюция дисциплины, сопровождаемая появлением новых материалов, устройств, а также введением новых концепций и характеристик;
- многофакторный характер описания объектов (НМ, устройств и т.д.) с привлечением большой совокупности параметров, определяющих структуру, состав, морфологию и другие особенности;
- взаимосвязь номенклатуры свойств и класса НМ, совмещение в номенклатуре молекулярных и макроскопических характеристик.

В результате, систематизация данных по наноматериалам намного превосходит по сложности классификации, сформированные в традиционных областях знания, например, в химии или материаловедении. Неоднократные попытки создания классификаторов для нанотехнологий [9] были зависимы от узкого

сегмента и ограниченной задачи (например, рубрикации Федерального портала (www.nanoportal.ru)) и, как правило, быстро отставали от текущего состояния предметной области. Для разрешения возникшей проблемы, существенно сдерживающей ее развитие, были предприняты международные усилия со стороны комиссии по численным данным (CODATA) и Международного научного совета (ICSU), создавших специальную рабочую группу с целью выработки единых рекомендаций, пригодных для широкого круга экспертов. Цели и принципы этой работы изложены в публикациях [10, 11]. Итогом работы стало появление в 2015 г. документа под названием Uniform Description System (UDS) [12], который по замыслу разработчиков, должен обеспечить универсальный подход к описанию как самих НМ, так и их свойств, технологий производства, сертификации, технической документации и т.д.. При характеристике наноматериалов система UDS призвана реализовать однозначность в его определении, т. е. надежно выделить один из объектов из множества близких ему по свойствам или подтвердить идентичность объектов по совокупности признаков: размеры и форма, структура, химическая природа, состояние поверхности, условия синтеза и т.п.

По сути, предложенный документ формирует именно систему метаданных, хотя и не обладает их формальными признаками, например в виде XML схемы или онтологии. Он оформлен в виде текстового документа для многоуровневого (и многоаспектного) анкетирования так, чтобы совокупность ответов на поставленные перед экспертом вопросы давала полное представление о НМ в соответствии с положенной в основу концептуальной схемой. Два момента отличают эту систему от аналогичных [13, 14]: максимальная широта охвата всей совокупности свойств и признаков; возможность добавления в систему новых рубрик, т. е. поддержка эволюции концептуальной схемы, что внушает надежду на решение большинства проблем систематизации данных.

В настоящей статье, рассмотрев в деталях особенности системы метаданных, мы постараемся ответить на три вопроса: 1) насколько полна и адекватна предложенная система; 2) в каких направлениях следует вести работу по детализации и расширению системы; 3) как привести систему к тому уровню формализации, который соответствует применяемым стандартам МД. Заметим, что авторы UDS сами рассматривают свою систему как возможную основу при формировании схемы БД или онтологии, избегая на данном этапе их реального проектирования.

ЛОГИЧЕСКАЯ СТРУКТУРА UDS И ФОРМАЛЬНАЯ ИДЕНТИФИКАЦИЯ НАНОМАТЕРИАЛОВ

Как и большинство систем метаданных, предлагаемое описание построено как иерархия соответствующих блоков информации, названных *категориями*. Четыре категории верхнего уровня (рис. 1) должны обеспечить максимальную полноту данных, необходимых для однозначного и исчерпывающего представления о НМ. Первая из них (**General Identi-**

fiers) дает формальное определение материала, включающее присвоение этому материалу названий и отнесение их к рубрикам выбранной классификационной схемы. Это простейший уровень идентификации, не затрагивающий как физических характеристик объекта, так и сведений о его производстве. Наиболее полную идентификацию дает вторая категория (**Characterization**), где в деталях перечисляются все позиции, требуемые для раскрытия сведений о свойствах: форма и размеры, химический состав, внутренняя и поверхностная структура и проч. Именно эта совокупность данных должна обеспечить однозначное выделение описываемого объекта из множества других с родственной структурой. Две последние категории (**Production** и **Specification**) охва-

тывают все моменты, связанные с производством и поставкой материала на рынок.

Каждая из указанных категорий порождает категории следующего уровня (или суб-категории). Так, первая категория **General Identifiers** (см. рис. 1) разбита на 4 суб-категории, соответствующие разным типам именования и классификации НМ, а категория **Characterization** – на три, соответственно его возможным типам. На следующем уровне введены категории, охватывающие физические характеристики материала: форму и размер, химический состав и проч. Набор этих характеристик (т. е., способ характеризации) зависит от типа материала, подтверждая взаимосвязь номенклатуры свойств и класса НМ, что отмечалось во Введении настоящей статьи.

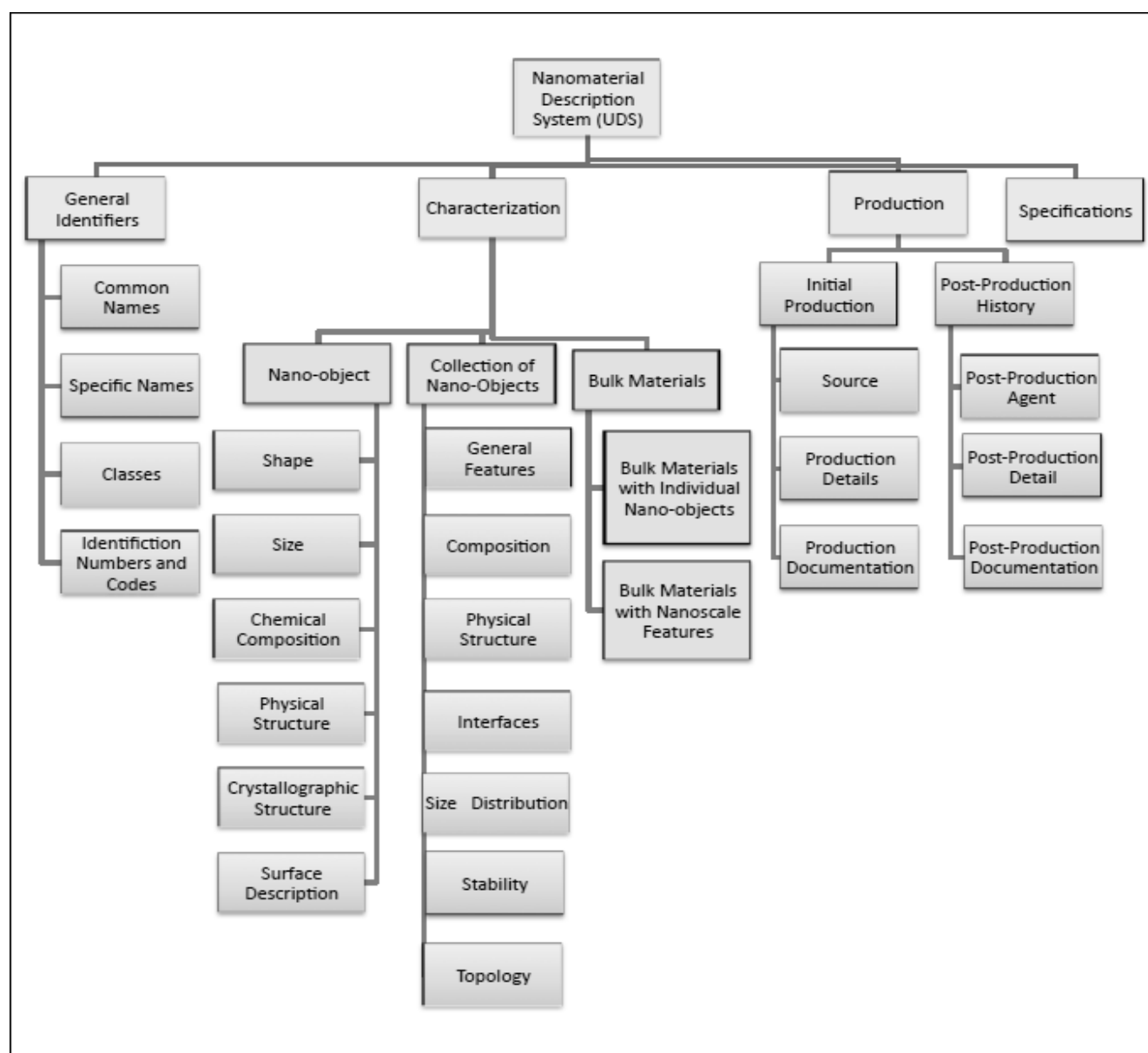


Рис. 1. Иерархия основных категорий в системе UDS [12]

Самый нижний уровень иерархии представлен так называемыми *дескрипторами*, т. е. метаданными, которые раскрывают содержание результата измерений или расчетов соответствующей характеристики, причем дескриптор определяет и формат представления данных: числовой (integer, float) или текстовой (string). Например, для характеристики формы нанобъекта (частицы, кластера, нанотрубки и т.п.) необходимы такие дескрипторы, как **число измерений на наномасштабе** или **число слоев**, которые задаются целыми числами, и такие, как **общее название формы** или **геометрическая регулярность**, предполагающие свободный текстовый формат. Разумеется, дескрипторы, относящиеся к категориям **Production** и **Specification**, соответствуют не физическим характеристикам, а технологиям или документам. Например, детали технологии раскрывают такие дескрипторы, как **число компонентов** (относится к категории **Исходные материалы**), **Общее описание** (относится к категории **Методы производства**) или **температура** и **давление** (относится к категории **Условия**). Первые два из указанных дескрипторов требуют данных в текстовом формате, вторые два – в числовом, с плавающей запятой.

Если иерархия категорий и суб-категорий структурирует предметную область, формируя ее концептуальную схему, то на уровне дескрипторов можно реализовать одно из важнейших требований к метаданным, обусловленных природой наноматериалов – необходимость перманентной подстройки, связанной с выбором их конкретного класса и поставленными задачами: фундаментальные исследования, разработка коммерческого продукта, технология производства, токсикология и экологическая оценка и др. Поэтому, как отмечают авторы [12], в общем случае не требуется использование всех дескрипторов, а их выбор зависит от потребности в определенных данных. Однако по мере появления новых материалов и методов их производства, а также новых характеристик и явлений может возникнуть необходимость в расширении системы МД за счет новых дескрипторов.

Типология наноматериалов. Как видно из рис. 1, структура метаданных связана с определенным соглашением (так называемое *нанодерево*), принятым ISO для систематизации всего мира наноструктур [15]. В этом соглашении основным понятием является НМ, который может быть в одной из двух форм: нанобъект и наноструктурированный материал.

К первой форме относятся объекты, у которых по крайней мере один размер соответствует наноразмерной шкале (до 100 нм), ко второй – макроскопические объекты, у которых наномасштабе соответствует внутренняя или поверхностная структура. Авторы [12] добавили к нанобъектам их ансамбли/коллекции, образованные намеренно или случайно. При этом, ансамбль соответствует тому же определению, что и нанобъект, т. е. имеет от одного до трех размеров в наномасштабе. Примером таких объектов служат так называемые «стручки» (peapods), образованные заполнением внутренней части углеродной нанотрубки (УНТ) несколькими молекулами фуллерена – C_{60} или C_{70} (рис. 2).

Еще одно изменение, внесенное в систематику НМ, относится к объемным материалам: считается, что их можно разбить на 2 субкласса (см. рис. 1). К первому субклассу отнесены материалы, содержащие идентифицируемые нанобъекты, ко второму – материалы, лишь проявляющие размерный эффект за счет наноразмерной внутренней или поверхностной структуры. Примером материалов первого субкласса являются нанокомпозиты (скажем, полимер с включениями в виде УНТ), примером второго – наноструктурированные стали или керамика.

Схема на рис. 1 фиксирует лишь три базовых типа наноматериалов, но не их дальнейшее разбиение на более узкие классы в соответствии с формой, составом и т.д., как было сделано в документе ISO [15]. Эту задачу решает категория **General Identifiers**.

Формальная идентификация наноматериала. Категория **General Identifiers** включает возможные названия НМ и отнесение его к определенной классификационной схеме. Скажем, популярному объекту (углеродные нанотрубки) достаточно присвоить одну из распространенных аббревиатур (УНТ) и отнести к рубрике **1.1.2 Линейные наноструктуры** научного рубрикатора Федерального портала «Нанотехнологии и наноматериалы» (www.portalnano.ru). Обе задачи (именование и классификация) в системе UDS решаются на двух уровнях: произвольном и согласованном с некоторыми стандартами. В качестве произвольного названия предлагается использовать любое англоязычное название, встречающееся в литературе, или их совокупность; для приведенного здесь примера с углеродными нанотрубками это могут быть CNT, SWCNT.

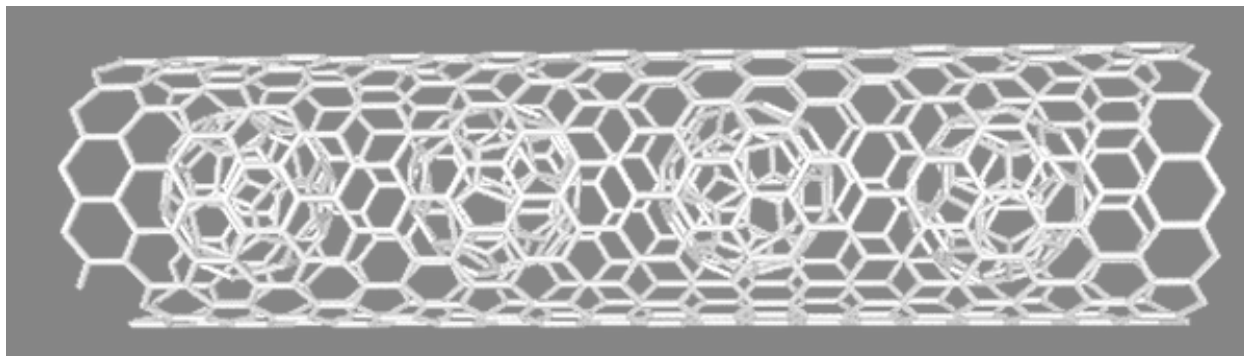


Рис. 2. Схематичный вид типичного ансамбля нанобъектов – структуры под названием «стручок» или peapod

Что же касается стандартизованного названия, то здесь, как и в химии, должны использоваться общедоступные классификаторы. В химии для идентификации вещества можно ссылаться на один из распространенных классификаторов, например регистр Chemical Abstract Service, регистры фундаментальных справочников, Beilstein для органических или Gmelin для неорганических веществ. Для наноматериалов возможности такой идентификации ограничены уже из-за достаточной «молодости» этой дисциплины. В частности, указанные регистры включают лишь те из наноструктур, которые можно определить посредством обычной химической формулы, например C_{60} или C_{70} .

Относительно широкий круг наноматериалов зарегистрирован в химической БД ChEBI [16], в которой они представлены аналогично обычным соединениям. Каждая из наноструктур определяется 5-значным индексом, который может играть роль стандартизованного названия. Например, индекс CHEBI:50796 определяет нанотрубки, а CHEBI:50853 – квантовые точки. Аналогичную функцию можно возложить на онтологию NPO [13], которая охватывает больший перечень НМ, хотя и с уклоном в сторону медико-биологических применений. Каждый объект в этой онтологии определен посредством ID вида NPO_***; например, NPO_606 определяет УНТ, а NPO_395 – наночастицу с внедренной в ее структуру молекулой лекарства. Недостатком обеих систем является то, что они совершенно не касаются объемных наноматериалов (композиты, порошки, нанопленки и проч.), которые лежат в основе большинства технологий и устройств.

Ограниченные возможности точного именования наноматериала компенсирует «встраивание» его в некоторую классификационную схему¹. Как и для названий, предполагается использование наряду с произвольной схемой, схемы, принятой на уровне международных стандартов. Произвольная классификация основывается, как правило, на одной из физических особенностей НМ: топологии, составе, проявлении квантовых эффектов и проч. Так, авторы [8, 9, 17] использовали детальную классификацию, основываясь на топологических признаках, когда за основу принимается число измерений $K = 0 \div 3$ макроскопического масштаба. Значение $K = 0$ соответствует кластеру, а $K = 3$ макроскопическому объекту, для которого приставка «нано» относится лишь к структурному элементу. Промежуточные значения $K = 1, 2$ соответствуют 1-мерным и 2-мерным структурам, у которых макроскопический масштаб охватывает одно или два измерения (нанопроволока или пленка). Если также аттестовать элементы, входящие в состав ансамбля нанообъектов или объемного материала, то можно получить простую «наноформулу» вида $KD\{L, M, N\}$, где индексы L, M, N относятся к составляющим элементам разного сорта, и выполняется очевидное неравенство $K \geq \max\{L, M, N, \dots\}$. Тем самым появляется возможность классифицировать не

только нанообъекты, но и их ансамбли, что предусмотрено схемой на рис. 1.

Применительно к углеродным наноматериалам неоднократно предлагалось классифицировать их по *гибридизации* химической связи – индексу sp^v , где целочисленные значения $v=1,2,3$ относятся к основным аллотропам углерода (карбин, графит, алмаз), а промежуточные, определяющие смешанные виды гибридизации, – охватывают множество углеродных нанотрубок [18].

В качестве стандартного классификатора авторы [12] рекомендуют так называемое «нанодерево», представленное в отчете ISO [15]. Распределение НМ по «ветвям дерева» проводится там по совокупности признаков, из которых на первом месте стоит топологический – число измерений в наномасштабе, которое изменяется от 1 до 3². При таком принципе классификации кластеры и наночастицы имеют размерность 3, нанопроволока или нанотрубка – 2 (2 наноразмера в сечении) и, наконец, пленки или графен – 1 (один наноразмер для толщины). В каждой из групп, определяемых размерностью, выделены три подгруппы: однокомпонентные, мультикомпонентные и наноструктурированные. Элементы этой иерархии, несколько измененные, предложены в UDS в качестве возможного стандартного классификатора. Изменения состоят лишь в том, что отмеченные три подгруппы вынесены на верхний уровень, как показано на рис. 1, а размерный признак применяется лишь к нанообъектам (см. табл. 1, где приведены характерные группы НМ).

Наряду с нанодеревом, рекомендованным в UDS в качестве признанного классификатора (*assigned by an authority* по терминологии [12]), могут быть использованы и другие классификаторы, например, приложение к PACS (Physics & Astronomy classification Scheme) (www.aip.org/publishing/pacs/nano-supplement), подробно описанное в работе [9]. Его достоинство в охвате большого числа НМ в соответствии с различными признаками, в том числе и сложных структур – ансамбли или нанопродукты (табл. 2).

Наконец, БД ChEBI [16], предложенную для именования наноматериалов, можно использовать и для отнесения их к определенной рубрике. Каждая из рубрик в зависимости от места, занимаемого в иерархии, может выполнять роль как идентификатора НМ, так и класса, охватывающего целую их группу, что иллюстрирует рис. 3, где показаны фрагменты иерархии для нанотрубок: в левой части рисунка код CHEBI:50796 определяет класс НМ, к которому относятся все наноматериалы, записанные ниже; их коды могут рассматриваться как стандартизованные названия. Некоторые из них, например CHEBI:50594 для carbon nanotube могут, однако, рассматриваться как идентификаторы класса, а «нижестоящие» – как идентификаторы НМ, например CHEBI:50595 для одностенных углеродных нанотрубок (single-walled carbon nanotube).

¹ Здесь имеется в виду классификатор самих наноматериалов в отличие от всей схемы на рис. 1, которая охватывает также их свойства и производственную историю.

² Разделение наноматериалов по числу измерений в наномасштабе не совпадает с общепринятым, когда за основу принимается число макроскопических измерений; в последнем случае кластер имеет размерность 0, нанопроволока – 1 и т.п. Именно эта классификация принята в работах [8, 9, 17].

Система наноматериалов, заимствованная в UDS из документа ISO [15]

Нанообъекты			Наноструктурированные материалы
3D	2D	1D	
Наночастица	Нановолокно	Нанопластина	Нанопорошок
Квантовая точка	Нанотрубка		Наноккомпозит
Онион	Наностержень		Нанопена твердая
	Нанопроволока		Нанопористый НМ
			Наножидкость
			Наноаэрозоль
			Наноструктурированная частица

Таблица 2

Набор рубрик, использованных в PACS для выделения сложноструктурированных наноструктур

PACS codes	Complex nanostructures
81.07.Pr	Organic-inorganic hybrid nanostructures
81.07.Lk	Nanocontacts
78.67.Pt	Multilayers; superlattices; photonic structures; metamaterials
78.67.Sc	Nanoaggregates; nanocomposites
78.67.Ve	Nanomicelles
63.22.Np	Layered systems
62.23.St	Complex nanostructures, including patterned or assembled structures

△ CHEBI:50796 nanotube △ CHEBI:50874 zinc oxide nanotube △ CHEBI:50875 molybdenum disulfide nanotube △ CHEBI:50876 silicon nanotube △ CHEBI:50878 aluminium nitride nanotube △ CHEBI:50879 tin disulfide nanotube △ CHEBI:50880 titanate nanotube △ CHEBI:50881 helical rosette nanotube △ CHEBI:52519 nitrogen-doped titanium dioxide nanotube △ CHEBI:50872 boron nitride nanotube △ CHEBI:50873 titanium dioxide nanotube △ CHEBI:50877 tungsten disulfide nanotube △ CHEBI:51115 copper(2+) sulfide nanotube △ CHEBI:50594 carbon nanotube △ CHEBI:50808 nitrogen-doped carbon nanotube △ CHEBI:50809 boron-doped carbon nanotube △ CHEBI:50810 boron-carbon-nitrogen nanotube	△ CHEBI:50594 carbon nanotube △ CHEBI:50595 single-walled carbon nanotube △ CHEBI:50596 multi-walled carbon nanotube ◇ CHEBI:50806 carbon nanotubosome ◇ CHEBI:50807 carbon nanorod ◇ CHEBI:50813 carbon nanorope △ CHEBI:50797 double-walled carbon nanotube
--	---

Рис. 3. Фрагменты иерархии наноматериала из БД ChEBI [16]

Ни одна из предложенных схем не способна охватить все множество наноматериалов, в том числе и вновь синтезированных. Поэтому столь важны предусмотренные в UDS возможности ее настройки на предметную область и потребности экспертного сообщества. В документе [12] декларируется, что НМ невозможно идентифицировать простым названием (или набором названий), что отличает их от обычных молекул. Решение этой проблемы, как ранее подчеркивалось в работах [8-11], лежит в привлечении максимального объема информации по свойствам и технологии производства, которую в систематизированном виде предоставляет UDS.

ИДЕНТИФИКАЦИЯ НАНОМАТЕРИАЛОВ ПО СОВОКУПНОСТИ ФИЗИЧЕСКИХ СВОЙСТВ

Основные идеи многофакторной характеристики НМ высказаны авторами UDS еще в докладе на рабочей группе, проведенной в 2012 г. в Париже под эгидой ICSU (Международного научного совета) и CODATA (комиссии по численным данным) [11]. Они исходили из факторов, определяющих ключевые отличия наноматериалов от традиционных (стали, сплавы, керамика и т.д.): большая величина отношения поверхность/объем; различия в электронной структуре поверхности и объема; квантовые размерные эффекты; «висящие» компоненты на поверхности; отличия химической реактивности при сравнении НМ с макроскопическим аналогом; синтез наночастиц, ранее не известных в макромире (УНТ, графен и т.п.); сильное влияние малых примесей; самосборка упорядоченных наноструктур; гетерогенный состав на малых масштабах.

Помимо отличий в свойствах наноматериалов, авторы [12] учли многообразие дисциплин, сталкивающихся с необходимостью детального описания НМ и решаемых при этом задач: проектирование устройств, предсказание свойств, токсикологические и экологические оценки, разработка стандартов и др. На этой основе были сформулированы требования к перечню определяющих факторов, причем этот перечень включал только те факторы, что необходимы для идентификации; за бортом оставлены многие из важных свойств НМ (термодинамические, электрические, оптические и т.д.), что позволяет сократить до разумного минимума весь объем метаданных. В соответствии с общей схемой МД (рис. 1), этот перечень различается для индивидуальных нанообъектов, их ансамблей (нанопродуктов) и макроскопических материалов.

Структура метаданных для индивидуальных нанообъектов. В табл. 3 представлены итоговые требования к наноматериалам в виде иерархии категорий, определяющих НМ по геометрии (форме и размеру), составу, физическим характеристикам и данным о поверхности. Отдельные аспекты каждой из характеристик раскрывают субкатегории, например, определяющие, с одной стороны, тип формы (цилиндр, сфера и т.п.), а с другой – ее особенности. Следующий (самый нижний) уровень иерархии занят дескрипторами, которые определяют элемент данных, задаваемый экспертом. Число дескрипторов достаточно велико, в соответствии с общим требованием многофакторного описания наноматериалов. При этом, каждый из дескрипторов снабжен в UDS точной *дефиницией*, определяющей содержание и формат вводимых данных. Для примера в табл. 4 приведены наборы дескрипторов, относящихся к форме и химическому составу НМ.

Таблица 3

Категории и субкатегории UDS, предложенные для характеристики индивидуальных нанообъектов

Shape	Size	Chemical composition	Physical structure	Crystallographic structure	Surface description
Shape Type	Applicable Dimensions	Atomic Composition	Layered Nano-Object	Physical Structure Identification	General Surface Description
Shape Features	Derived Dimensions	Molecular Composition	Shell Structure	Unit Cell Information	Surface Treatment
	Internal Dimensions	Chemical Moieties	Physical Features	Basic Unit Cell Parameters	Surface Geometry
			Defects		Surface Electronic Properties
			Entrapment		Other Surface Properties
			Additions		

Дескрипторы, определяющие форму и химический состав наноматериалов

Shape		Chemical Composition		
Shape Type	Shape Features	Atomic composition	Molecular composition	Chemical Moieties
Number of dimensions on nanoscale	Type of feature	Number of different types of atom present	Number of different types of molecule present	
General shape	Regularity of feature	Atom present	Molecules present	
ISO 229 shape name	Number of feature	Composition percentage type	Molecular formula	
Specific shape	Symmetry of feature	Composition percentage	Molecular name	
Type of thickness of a nano-object with one dimension at the nanoscale		Type of composition	Structural formula	
Cross-sectional view for nano-object with two dimensions at the nanoscale		How measured	CAS Registry number	
Number of layers (for nano-object with two or three dimensions at the nanoscale)			IUPAC InChI	
Geometric regularity			Composition percentage type	
Shape symmetry			Composition percentage	
Symmetry components			Type of composition	
			How measured	

Примечание: перечень дескрипторов в субкатегории **Chemical Moieties** (часть или функциональная группа молекулы) идентичен тому, что приведен в субкатегории **Molecular Composition**.

Помимо двух основных форматов (текстового и числового), в UDS используется (хотя специально не оговаривается) формат, именуемый *enumeration*, когда элемент данных принимается из ограниченного списка или внешнего контролируемого словаря. Так, дескриптор **ISO 229 Shape name** (табл. 4) определяет форму в соответствии с принятым в стандарте ISO [16] списком из пяти возможных форм: nanoparticle, nanorod, nanotube, nanoplate, nanoscone. Аналогично, при определении состава нано-объекта названия и обозначения атомов задаются в соответствии с таблицей Менделеева, а молекулу однозначно идентифицируют номером из регистра CAS. Другие дескрипторы, использующие этот формат (в табл. 4 они закрашены темным цветом), позволяют путем выбора из списка определить: процентное содержание компонента (по массе или числу частиц), «тип состава» (отдельное измерение, вычисленное значение или результат усреднения), метод измерения (дескриптор *how measured*). Очень эффективным оказывается этот формат при детализации категории **Physical Structure**. Например, чтобы раскрыть содержание **Physical Features**, предложен их априорно состав-

ленный перечень (отверстие, выпуклость/выступ, отросток, наконечник, ножка и др.), позволяющий отразить морфологию нанообъекта.

Формат *enumeration* широко использовался в ThermoML [5] – системе метаданных для представления теплофизических свойств веществ, позволяя фиксировать тип свойства, фазовое состояние, методы измерения и очистки, статус данных и вид их неопределенности и проч. Этот формат исключает произвол в записи данных, открывая богатые возможности в машинной обработке при поиске и обосновании логического вывода. Заметим, что текстовый формат (string) также не всегда предполагает свободное описание, но может следовать строгим правилам, как, например, при записи молекулярной формулы, названия и структурной формулы, которые следуют принятому стандарту одной из общедоступных химических БД, например ChEBI [16] или ChemSpider (www.chemspider.com). Дескриптор **IUPAC InChI** содержит особый тип химических данных – так называемую линейную нотацию в виде строки символов, кодирующих сведения о составе и топологии молекулы.

Полный перечень всех дескрипторов представлен в тексте руководства UDS на сайте IUPAC [12]). Расширяя или изменяя перечень дескрипторов, можно провести *тонкую настройку* структуры метаданных на требования, обусловленные спецификой объектов или решаемых задач.

Структура метаданных для ансамбля нанообъектов. Как видно из диаграммы на рис. 4, в структуре МД можно выделить три основных блока: состав ансамбля, описание каждого из его компонентов, описание ансамбля в целом. Состав задается примерно так же, как и состав молекулы: указываются типы индивидуальных нанообъектов и их число в ансамбле. Например, для структуры на рис. 2 («стручок») описание должно указывать на два типа нанообъектов (УНТ и фуллерен) в количестве одной УНТ и 4-х молекул фуллерена. Следующий блок должен описать каждый из нанообъектов в соответствии с предусмотренной структурой (табл. 3), дополняя их информацией по производственному циклу. Наконец, пять категорий (**Physical Structure, Interfaces, Size distribution, Stability, Topology**) представляют различные аспекты ансамбля как целого. Первая из этих категорий дает общие характеристики: форма и размеры, взаимоположение объектов, их ассоциация и т. п. Категория **Interfaces** включает описание поверх-

ностей, разделяющих индивидуальные нанообъекты, их форму, размерность, регулярность и проч., причем большинство дескрипторов предусматривают лишь качественную характеристику. Напротив, дескрипторы категории **Size distribution**, в основном, предполагают численную характеристику, а метод измерения и среда могут быть элементом классификатора (формат *enumeration*).

Категория **Stability** определяет устойчивость ансамбля по отношению к спонтанному распаду или внешнему воздействию: тепловому, химическому, и др. Сначала субкатегория **Overview** задает тип нестабильности из некоторого перечня и дает ответ на вопрос об ожидаемости распада. Последующие субкатегории в деталях раскрывают эти виды нестабильности.

Наконец, последняя категория (**Topology**), в трактовке авторов [12], характеризует связность и непрерывность ансамбля, проявляющиеся в их влиянии на свойства. Пример дает молекулярная биология, где функциональность объекта оказывается зависящей от *гомеоморфизма*, т. е. способности объекта к непрерывной трансформации в другой без разрывов и связывания. В целом, топология ансамбля достаточно тонкое понятие, которое может быть раскрыто лишь на качественном уровне, без разбиения на уровни иерархии.

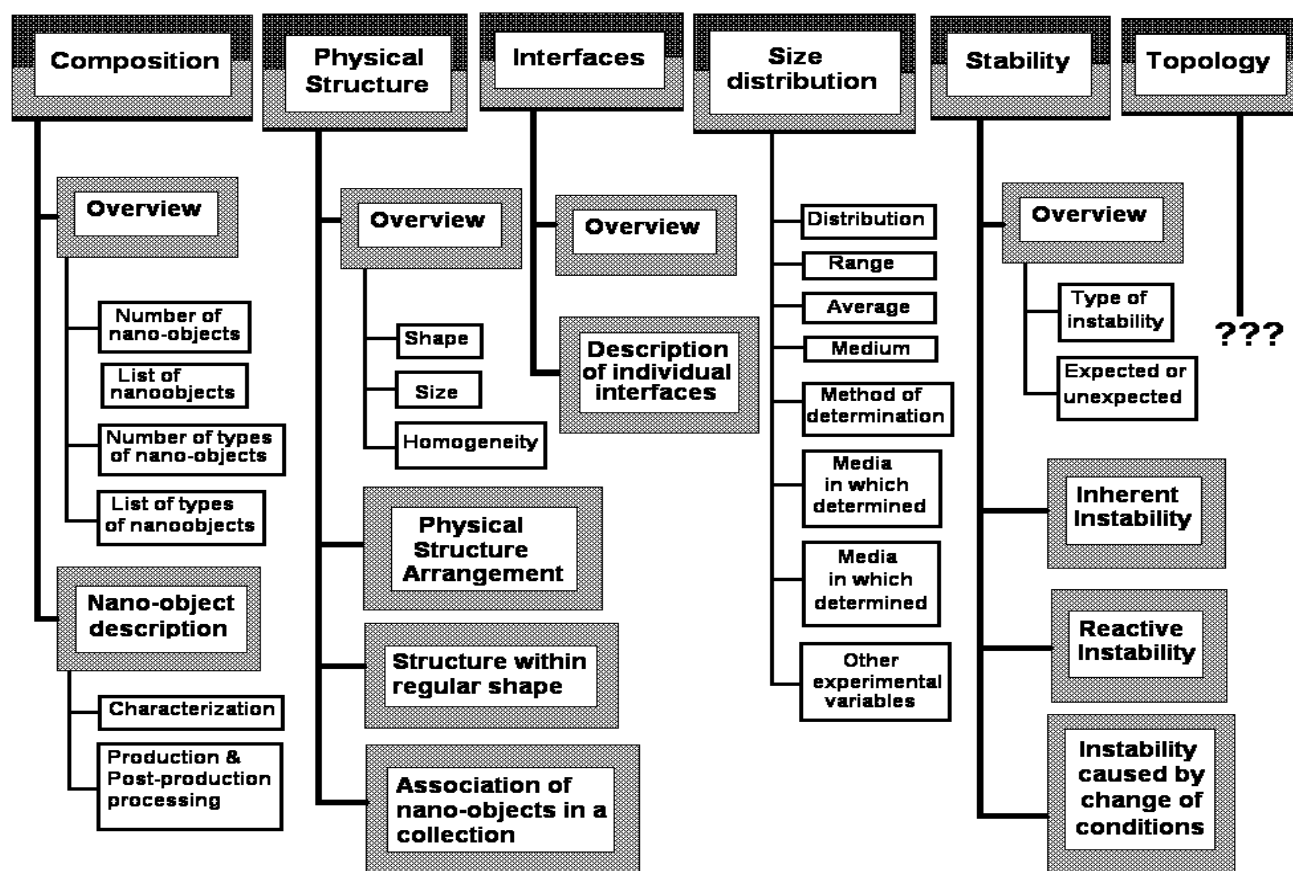


Рис. 4. Иерархия категорий и дескрипторов, определяющих ансамбль нанообъектов

Объемные (макроскопические) наноматериалы. Применительно к этому классу НМ специальной иерархии категорий не предложено. Для двух возможных типов объемных НМ, отмеченных на рис. 1, возможности системы UDS различны. Если объемный НМ включает идентифицируемые нанообъекты, то важную роль играет степень свободы, которую эти нанообъекты имеют в пределах соответствующей фазы. Как в жидкой фазе, так и в твердой свободное движение позволяет ограничиться теми предписаниями, которые предложены для самого нанообъекта, либо его ансамбля. Если же нанообъект (скажем, металлическая наночастица) жестко закреплен в твердой матрице, то его описание должно быть дополнительно описанием самой матрицы, следуя тем стандартам, что предложены для типовых материалов: металлов, керамики, полимеров и т.д. Например, для нанокompозита, образованного включением УНТ в полимерную матрицу, потребуется детализировать оба материала. Разумеется, возможны и промежуточные случаи ограниченной свободы нанообъекта в пределах матрицы, которые потребуют расширения возможностей UDS.

Второй из двух отмеченных типов на рис. 1 (*Bulk Materials with Nanoscale Features*) охватывает наноструктурированные материалы (стали, керамика и проч.), элементы которых (кристаллиты) попадают в наномасштаб, что заметно сказывается на физических и эксплуатационных свойствах. Материалы этого типа уже сейчас нашли широкое применение в машиностроении, энергетике, атомной технике и др. отраслях [19, 20]. Текущая версия UDS, к сожалению, не дает возможностей для идентификации объемных НМ. В последнем разделе настоящей статьи мы обсудим возможности их описания, основываясь на доступных онтологиях из области материаловедения.

МЕТАДААННЫЕ ДЛЯ ПРОИЗВОДСТВЕННОЙ ИСТОРИИ И УСЛОВИЙ ПОСТАВКИ

Две последние категории верхнего уровня иерархии на рис. 1 (**Production** и **Specification**) относятся к производству и аттестации наноматериала как коммерческого продукта. Ключевая особенность НМ – сильная зависимость свойств от технологии изготовления вынуждает отражать в системе метаданных неразрывную связь свойств и условий производства. Полноту информации в UDS обеспечивает разбиение категории **Production** на две, отражающие начальную стадию и стадии, следующие за производственной (см. рис. 1). МД для начальной стадии охватывают состав и количество исходных материалов, метод производства и рецептуру, оборудование, характер и состояние среды, данные по целевому продукту. Аналогичная структура МД в категории **Post-production history** отражает последующие стадии жизненного цикла: очистку, хранение, транспортировку и т.п. Наиболее значима результирующая группа дескрипторов, своеобразный «паспорт» на поступающий потребителю продукт (табл. 5).

Как только наноматериал становится коммерческим продуктом, возникает необходимость в массиве сопровождающих документов, объединенных по-

нятием «спецификация», неотъемлемой составляющей любой продукции – от сырья до высокотехнологичного оборудования. Для каждой из отраслей науки и производства требования к составу и содержанию спецификации регламентируются национальными или международными стандартами. Применительно к НМ руководящим документом является стандарт ISO [21], который определяет необходимый объем данных по свойствам, условиям хранения и поставки, контрольным измерениям и проч. Что же касается категории **Specification** в системе UDS, то она образует остов метаданных, определяющих формальные сведения: название, версию и номер спецификации, дату ее выпуска и официальный статус, выпускающую организацию.

Таблица 5

Итоговая характеристика коммерческого продукта

Post-Production Process Result
Nanomaterial produced
Purity
Composition
Yield
Physical state
Date produced
Location
Post-producing organization
Batch number
Post-producing documentation

Надо заметить, что иерархия категорий, описывающих производственную историю и ссылающихся на спецификацию, почти одинакова для многих отраслей, особенно связанных с производством материалов³. Поэтому для этих целей можно использовать ряд словарей и онтологий общетехнического профиля, например, стандарт ISO 10303-235: “**Engineering properties for product design and verification**” [22]. Это позволит несколько снизить нагрузку на UDS, перенося ссылки на другие онтологии и словари.

ПОТЕНЦИАЛ СОВЕРШЕНСТВОВАНИЯ СИСТЕМЫ UDS

Чрезвычайная сложность поставленной разработчиками системы задачи – однозначная и исчерпывающая идентификация всех возможных типов наноматериалов, естественно, не позволяет рассматривать предлагаемую версию как окончательную. Помимо многообразия объектов, проблему в стандартизацию описания вносит многообразие требований, возникающих перед исследователями, разработчиками продуктов, потребителями, создателями стандартов, медицинских и экологических норм, специалистов по

³ Имеется в виду близость информационных моделей для различных отраслей, но никак не содержание спецификаций, различающихся для каждой отрасли и вида продукции.

информатике и др. В самом документе UDS [12] неоднократно подчеркивается, что по мере накопления знаний, могут потребоваться новые дескрипторы, т. е. новые позиции в структуре метаданных. Более того, для некоторых НМ, по мнению авторов [12], практически отсутствует научно обоснованная база для стандартизации наноструктурированных объемных материалов (вторая группа *Bulk Materials*, указанная на рис. 1) или некоторых характеристик ансамблей, например, топологии (см. рис. 4). Поэтому созданная система МД стимулирует работы по ее совершенствованию и адаптации к отдельным сегментам нанотехнологии. При этом возможны два направления такой деятельности: (1) ревизия и расширение логической структуры, включая вопросы классификации, выбора форматов, согласования с контролируемыми словарями и онтологиями; (2) расширение тематики путем введения дополнительных дескрипторов.

Первое направление – необходимое условие при переходе к строго формализованным онтологиям или БД, что рассматривается в UDS как одно из важнейших применений разработанной системы⁴. Некоторые из требуемых здесь шагов обсуждались в ходе подготовки настоящей статьи. Среди них максимально возможное использование в дескрипторах формата *enumeration* взамен свободного текста, что является условием машинной реализации поиска и логического вывода. К этому типу данных можно обращаться при определении химического состава, вида физических и морфологических особенностей, методов измерения и проч.

Второе направление – широкое использование ссылок на внешние ресурсы в виде контролируемых словарей, например, упомянутой БД ChEBI [16] или онтологии NPO [13] для именования и систематизации нанообъектов. Внешние ресурсы можно использовать и как терминологические источники, например, для характеристики типов поверхности или элементов кристаллической структуры. В качестве возможных ресурсов здесь можно указать на компендиум химической терминологии **Goldbook** [2] или международный стандарт кристаллографических данных CIF (www.iucr.org/resources/cif). Активное привлечение словарей и онтологий, размещенных в Интернете, рассматривается сейчас как одно из важнейших условий глобальной интеграции данных [3].

Наиболее трудная задача в совершенствовании UDS – идентификация объемных материалов, для которых разработчики системы отказались от построения таксономии из категорий и дескрипторов. Указанный выше путь привлечения внешних ресурсов позволяет подойти естественным образом к ее реше-

нию, одновременно расширяя и логическую структуру метаданных, и предметную область. Соответственно, двум типам объемных наноматериалов требуется привлечение разных ресурсов. К первому типу НМ относятся объемные материалы, содержащие идентифицируемые нанообъекты – типичными представителями таких материалов являются нанокompозиты, например, полимеры с наноразмерными включениями. Для этого класса НМ уже разработана соответствующая онтология [14], основу которой составляет таксономия классов и экземпляров, включающая такие классы, как **MaterialType**, **NanoObject**, **Nanocomposite**, а также классы, объединяющие химические характеристики, такие, как **ChemicalIdentity**, **ChainComposition**, **Structure**. Класс **MaterialType** включает 5 дочерних классов (**CarbonBased**, **Ceramic**, **Metallic**, **Polymer**, **SiliconBased**), что в сочетании с довольно простой таксономией нанообъектов (**NanoFiber**, **NanoFilm**, **NanoLayer**, **NanoPowder**, **NanoSurface**) позволяет отразить формирование разных типов нанокompозита, определяемого выбором матрицы и наполнителя. При построении онтологии [14] использовался тот же принцип, что здесь рекомендуется для расширения UDS: заимствовался класс **Polymer** (совместно с его дочерними классами) из онтологии **ChemAxiom** и класс **Metallic** из онтологии **PeriodicTable**.

Однако группа наноматериалов, содержащих идентифицируемые нанообъекты, не сводится только к композитным материалам. Можно назвать немало материалов, построенных из нанообъектов без выделения матрицы и наполнителя, как в случае композита. Примером такого материала является фуллерит, молекулярный кристалл, образованный из молекул C_{60} или C_{70} за счет Ван-дер-Ваальсова взаимодействия [23]. Такой кристалл может существовать также в виде одномерных или двумерных полимеров. Многообразные пленки, волокна, образцы «нанопружи» и «нанобумаги» производятся из ориентированных углеродных нанотрубок [24]. Красивый пример подобного наноматериала дает макроскопический ультра-легкий аэрогель с плотностью порядка mg/cm^3 [25]. Такой аэрогель имеет вид трехмерной углеродной структуры, которая построена из графеновых листов, связанных ребрами из УНТ. Все такие материалы по своей природе относятся к ансамблю нанообъектов с тем условием, что в этом классе будет предусмотрено наличие макроскопических объектов. Для этого перечни классов и свойств, используемых для их характеристики, должны быть расширены так, чтобы можно было представить такие типичные макроскопические свойства, как, например, плотность, модуль Юнга, теплоемкость и т.п.

В отношении второй группы наноматериалов (см. рис. 1), которые не содержат идентифицируемых нанообъектов, текущая версия UDS не дает никаких рекомендаций по их описанию. По сути, однако, эти материалы могут быть отнесены к классам традиционных конструкционных материалов, что открывает интересную возможность использования онтологий, развитых в материаловедении для обмена данными по физическим и эксплуатационным свойствам. К наноматериалам относятся те конструкционные ма-

⁴ Фраза из раздела “Use of the Uniform Description System” [12]: «The UDS provides a backbone for building the database schemas and ontologies that are at the core of a nanoinformatics resource so that information from different resources can be compared and contrasted correctly» (Руководство UDS обеспечивает основу при построении БД и онтологий в области наноинформатики, что позволяет сравнивать и проверять на достоверность данные из различных источников – перевод автора).

териалы, которые за счет специальных технологических приемов приобрели наноразмерные особенности, т. е. заметную зависимость свойств от размера структурной единицы [19, 20]. Поэтому в онтологии придется ввести новый параметр, дополняя обычные параметры, определяющие состояние среды (температура, давление и т.п.). Более того, при характеристике НМ могут понадобиться более тонкие характеристики, например, распределение структурных единиц по размерам. Для характеристики подобных материалов, включая их механические, тепловые и эксплуатационные свойства, может быть привлечена одна из онтологий, развитых в рамках материаловедения, например [26, 27]. Однако, привлекая онтологию по конструкционным материалам, необходимо предусмотреть необходимость расширения набора параметров состояния.

Таким образом, очевидный путь расширения системы метаданных состоит в ее максимально возможной интеграции с уже существующими онтологиями – как с теми, что отражают некоторые классы НМ, так и с онтологиями общего содержания, относящимися к химии, материаловедению и другим дисциплинам. Эта же стратегия, как отмечалось, может применяться для отражения не только физических характеристик наноматериалов, но и жизненного цикла изделия.

СПИСОК ЛИТЕРАТУРЫ

- Бездушный А.Н., Вежнев А.А., Серебряков В.А., Шкотин А.В. Метаданные: определение и использование. – URL: www.elbib.ru/index.phtml?page=elbib/rus/methodology/metadata
- IUPAC Compendium of Chemical Terminology - the Gold Book. – URL: <http://goldbook.iupac.org/>
- Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Серебряков В.А., Теймуразов К.Б. Технология научных публикаций в среде «Открытых связанных данных» // Научно-техническая информация. Сер. 1. – 2013. – №12. – С. 1-11.
- Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А. Роль метаданных в создании и использовании информационных ресурсов о свойствах веществ и материалов // Научно-техническая информация. Сер. 1. – 2008. – №11. – С. 13-19.
- Frenkel M. Global communications and expert systems in thermodynamics: Connecting property measurement and chemical process design // Pure Applied Chem. – 2005. – Vol. 77, № 8. – P. 1349-1367.
- ThermoML Schema. – URL: <http://media.iupac.org/namespaces/ThermoML/ThermoML.xsd>
- Kaufman J.G., Begley E.F. A Data Interchange Markup Language // Advanced Materials & Processes. – November 2003. – P. 35-36.
- Елецкий А.В., Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Трахтенгерц М.С. Теплофизические свойства наноразмерных объектов: систематизация и оценка достоверности данных // Теплофизика высоких температур. – 2012. – Т. 50, № 4. – С. 524-532.
- Буйлова Н.М., Елецкий А.В., Зицерман В.Ю., Кобзев Г.А. Систематизация источников и данных по нанотехнологиям // Научно-техническая информация. Сер. 1. – 2013. – №11. – С. 31-42.
- Rumble J., Freiman S. Describing nanomaterials: meeting the needs of diverse data communities // Data Science Journal. – 2012. – Vol.11. – P. 1 – AMSD – 6.
- Rumble J., Freiman S., Teague C. The Description of Materials // ICSU-CODATA Workshop on Nanomaterials February 2012.
- Uniform Description System for Materials On the Nanoscale. Prepared by the CODATA-VAMAS Working Group On the Description of Nanomaterials. – URL: www.codata.org/nanomaterials (Version 1.0, 1 February 2015).
- NPO: NanoParticle Ontology for Cancer Nanotechnology Research. – URL: www.nano-ontology.org
- Карпова И., Порысева Е., Казаков Г., Кольцова Э. Разработка онтологии в области наноконпозиционных материалов // Информационные ресурсы России. – 2012. – №2. – С. 5-9.
- ISO/TR 11360:2010, Nanotechnologies – Methodology for the classification and categorization of nanomaterials. ISO technical committee ISO/TC 229, Nanotechnologies. Retrieved from the World Wide Web, October 20, 2012. – URL: www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=55967
- Degtyarenko K., de Matos P., Ennis M. et al. ChEBI: a database and ontology for chemical entities of biological interest // Nucleic Acids Research. – 2008. – Vol. 36. - Data base issue. – D344–D350.
- Pokropivny V.V., Skorokhod V.V. New dimensionality classifications of nanostructures // Physica E. – 2008. – Vol. 40, № 7. – P. 2521-2525.
- Heimann R.B., Evsyukov S.E., Koga Y. Carbon allotropes: a suggested classification scheme based on valence orbital hybridization // Carbon. – 1997. – Vol. 35. – P. 1654-1658.
- Андреевский Р.А. Основы наноструктурного материаловедения: возможности и проблемы. – М.: Бином. Лаборатория знаний, 2012. – 251 с.
- Гусев А.И. Наноматериалы, наноструктуры, нанотехнологии. 2-е изд. – М.: Физматлит, 2007. – 416 с.
- ISO/TS 12805:2011 (en). Nanotechnologies-Material Specifications-Guidance on specifying nano-objects. – URL: www.iso.org/obp/ui/#iso:std:51766:en
- Swindells N. The representation and exchange of material and other engineering properties // Data Science Journal. – 2009. – Vol. 8. – P. 190-200.
- Маркин А.В. Термодинамика кристаллических полимерных наноструктур фуллерена C60 / Дисс.... канд. хим. наук. – Н.Новгород: Нижегородский госуниверситет им. Н.И. Лобачевского. – 2004. – 213 с.
- Елецкий А.В. Углеродные нанотрубки // УФН. – 1997. – Т. 167, №9. – С. 945-972.

25. Лаверов И. Аэрогель из графена и углеродных нанотрубок лишен недостатков своих предшественников // Сайт ЭЛЕМЕНТЫ. – URL: <http://elementy.ru/news/432045>
26. Cheung K., Hunter J., Drennan J. MatSeek: An Ontology-Based Federated Search Interface for Materials Scientists // IEEE Intelligent Systems. – 2009. – Vol. 24, № 01. – P. 47-56.
27. Ashino T., Fujita M. Definition of a web ontology for design-oriented material selection // Data Science Journal. – 2006. – Vol. 5. – P. 52-63.

Материал поступил в редакцию 26.05.15.

Сведения об авторах

ЕРКИМБАЕВ Адильбек Омирбекович – кандидат технических наук, зав. лабораторией баз данных

Объединенного Института высоких температур РАН (ОИВТ РАН), Москва
e-mail: adilbek@ihed.ras.ru

ЗИЦЕРМАН Владимир Юрьевич – кандидат физико-математических наук, ведущий научный сотрудник лаборатории баз данных ОИВТ РАН
e-mail: vz1941@mail.ru

КОБЗЕВ Георгий Анатольевич – доктор физико-математических наук, главный научный сотрудник, советник НИЦ электрофизики и тепловых процессов ОИВТ РАН
e-mail: gkbz@mail.ru

ТРАХТЕНГЕРЦ Михаил Самойлович – кандидат технических наук, ведущий научный сотрудник лаборатории баз данных ОИВТ РАН
e-mail: trachtengerts@yahoo.com

Л.В. Шемберко, А.И.Слива

Библиографическая база данных по социологии: принципы и инструменты формирования, структура, особенности поиска

Описываются цели и задачи создания библиографической базы данных по социологии в Институте научной информации по общественным наукам (ИНИОН) РАН. Раскрывается ее проблемно-тематическая структура, видовой и языковой состав документов. Рассматриваются основные принципы аналитико-синтетической переработки документов и запросов на основе комплекса лингвистических средств по социологии. Отмечается необходимость междисциплинарного доступа к информационным ресурсам (отечественным и зарубежным) для повышения качества научно-информационного обеспечения потребителей социологической информации.

Ключевые слова: социологическая информация, информационные потребности, базы данных, информационно-поисковые тезаурусы, электронные ресурсы, стратегии поиска

ВВЕДЕНИЕ

Повышение эффективности информационного обеспечения социологических исследований неразрывно связано с широким распространением новых информационных технологий и оперативным доступом потребителей к отечественным и зарубежным информационным ресурсам в области социологии. Как известно, термин «социология», который в буквальном смысле означает «наука об обществе», ввел в оборот в первой половине XIX в. французский ученый Огюст Конт (Auguste Conte). В широком смысле социология – это система знаний об общественном устройстве жизни людей, о природе и разнообразных проявлениях человеческой общности. В узком смысле социология занимается анализом структуры социальных отношений, процессов социального взаимодействия, группового поведения.

Социология развивалась в основном в Западной Европе, но свой вклад в социологическую науку внесли и русские мыслители: Г. В. Плеханов, П. Л. Лавров, Н. К. Михайловский, В. И. Ленин, Н. Я. Данилевский, П. А. Сорокин. Одновременно с так называемой академической социологией в России большое развитие получила и политическая социология. Испытывая влияние различных течений западной социологии, российская социология вместе с тем смогла выдвинуть немало собственных теорий и концепций, отразивших своеобразие развития российского общества и связанных с именами таких русских мыслителей, как А. Хомяков, Н. Бердяев, К. Леонтьев, Вл. Соловьев и др.

Начало возрождения отечественной социологии приходится на середину 50-х гг. XX в. В 1958 г. была создана Советская социологическая ассоциация, в рамках которой в 1989 г. организовано Российское

общество социологов (РОС), в составе которого насчитывается свыше 70 региональных отделений. В 1987 г. был создан Всесоюзный центр изучения общественного мнения (ВЦИОМ), а также ряд независимых социологических служб. Значительный вклад в развитие современной социологической мысли внесли известные отечественные социологи: И.В. Бестужев-Лада, А.Г. Здравомыслов, Ю.А. Левада, Т.И. Заславская, Н.А. Аитов, Ж.Т. Тощенко, С.С. Фролов, Г.А. Осипов, В.А. Ядов, В.И. Добреньков и др.

В настоящее время наблюдается активный процесс развития социологии как самостоятельной научной дисциплины. При этом социология как наука о закономерностях становления и развития социальных систем, общностей, групп, личностей подразделяется на множество исследовательских областей, отраслевых направлений, тематических областей и сфер, которые выделились в процессе дифференциации социологического знания. Необходимо особо подчеркнуть, что социология призвана не просто выявлять и описывать тенденции общественного развития, но и объяснять, интерпретировать их, предлагать варианты практического решения тех или иных проблем.

Как фундаментальная наука, социология объясняет социальные явления, собирает и обобщает информацию о них. Как прикладная наука, социология позволяет прогнозировать социальные явления и управлять ими. При этом важные сведения, которые накопила социология об общественных явлениях и процессах, используются в социальном управлении для воздействия на экономическое, политическое и правовое поведение человека. Большую роль в современном обществе играет изучение общественного мнения: взглядов людей, их ценностных ориентаций, представлений о социальных проблемах образования,

семьи, здоровья, о растущей угрозе наркотизации и алкоголизации и т. д. Социология стремится познать, как осуществляется переход от научных дисциплин к небольшим интеллектуальным образованиям, что позволяет поддерживать дисциплинарную идентичность, а также в чем заключается социальный механизм создания успешных теоретических групп.

ОСНОВНЫЕ КАТЕГОРИИ ПОТРЕБИТЕЛЕЙ СОЦИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ

Место социологии среди других социальных наук, предметы ее исследования, значение для изучения всех сфер общественного развития, определяют основной круг потребителей социологической информации. К ним относятся ученые, которые занимаются фундаментальными и прикладными исследованиями в различных областях социологического знания, преподаватели вузов, студенты (бакалавры, магистры), специалисты из смежных областей (философы и обществоведы, культурологи и искусствоведы), специалисты из сферы образования (педагогика и психологии). Особую группу составляют потребители социологической информации, работающие в аналитических социальных центрах, PR-службах, консалтинговых компаниях, различных СМИ и издательствах, рекламных компаниях, сотрудники государственных и муниципальных органов власти, представители политических партий, изучающие мнение электората в предвыборных кампаниях, а также кадровые службы, маркетинговые отделы предприятий и фирмы, представляющие практические направления социологии.

Среди коллективных потребителей информации по социологии следует назвать, прежде всего, Институт социологии РАН, который осуществляет разработку фундаментальных проблем социологической теории, проводит масштабные эмпирические исследования в области современных социальных и политических практик, а также принятия эффективных управленческих решений. В структуре Института двенадцать исследовательских центров, Центр социологического и политологического образования, и три филиала (Башкирский в г. Уфа, Приволжский в г. Нижний Новгород, Южнороссийский в г. Ростов-на-Дону). Основные направления научных исследований этих организаций во многом зависят от информации по истории социологии и общей социологической теории, большой интерес представляют методологии и методы социологических и политологических исследований, а также социологическая информация по трансформации социальной структуры российского общества и социальной мобильности; трансформации социальных институтов в процессе формирования гражданского общества; данные о динамике изменения ценностей российского общества, особенностях формирования новых идентичностей; социологии управления; трансформации семьи и гендерных отношений; социологии политических отношений и электоральных отношений.

В 1989 г. был создан Санкт-Петербургский филиал Института социологии Академии наук СССР, преобразованный позднее в Социологический институт РАН. Информационные потребности сотрудников

данного Института связаны с разработкой теоретико-методологических и инструментальных средств описания, объяснения и прогнозирования социальных процессов, с проведением сравнительных научных исследований в области истории социологии и других социальных наук в России, с изучением трансформации социальной структуры и изменений массового сознания в современной России, социально-институциональной динамики властных групп и общественных движений, а также анализом социальных механизмов в научной, инновационной деятельности и в социально-трудовой сфере.

Потребителями социологической информации являются также сотрудники Института социологии образования Российской академии образования и Независимого института социальной политики, учредителями которого являются Всероссийский центр изучения общественного мнения (ВЦИОМ) – старейший в России социологический центр и Российская Академия народного хозяйства и государственной службы при Президенте РФ. В социологии образования особый интерес исследователей привлекают проблемы коммерциализации высшей школы, переход многих отечественных вузов на платную форму обучения, введение единого государственного экзамена, элементы рыночных отношений в школе и др. К каким последствиям приведут эти изменения, как они скажутся на социальной структуре общества, не углубят ли социальное неравенство в стране, не станут ли противоречия в системе образования серьезным тормозом общественного развития страны – в поиске ответов на эти вопросы социологи принимают самое непосредственное участие.

В последние годы в российском обществе стал остро ощущаться социальный запрос на общественное, политическое и научное исследование границ, ответственности, механизмов и результатов социальной политики. При этом одной из наиболее сложных задач совершенствования социальной политики в стране является установление диалога между государством и обществом. Важную роль в решении этой задачи начинают играть информационные ресурсы, банки социологических данных, базы социологической информации, а также отечественные и зарубежные электронные журналы, например, такие, как электронный журнал по социальной политике SPERO – Social Policy: Expertise, Recommendations, Overviews (<http://spero.socpol.ru/>).

В настоящее время подготовку социологов осуществляют более сотни вузов России, профессорско-преподавательский состав которых заинтересован в получении отечественной и зарубежной информации по социологии для составления учебных программ и проведения научных исследований. Среди них следует назвать, прежде всего, следующие: Московский государственный университет им. М.В.Ломоносова (МГУ), факультет социологии; Московский государственный институт международных отношений МИД России (МГИМО), факультет международной журналистики (кафедра социологии); Государственный академический университет гуманитарных наук (ГАУГН), институт социологии; Государственный университет Высшая школа экономики (ГУ-ВШЭ),

факультет социологии; Государственный университет управления (ГУУ), институт психологии и социологии управления; Московский государственный лингвистический университет (МГЛУ), Институт международных отношений и социально-политических наук (кафедра социологии); Московский государственный педагогический университет (МГПУ), Институт психологии, социологии и социальных отношений; Московский педагогический государственный университет (МПГУ), факультет социологии, экономики и права; Российская академия государственной службы при Президенте Российской Федерации (РАГС), кафедра социологии; Российский государственный гуманитарный университет (РГГУ), социологический факультет; Российский государственный социальный университет (РГСУ), факультет социологии; Российский государственный торгово-экономический университет (РГТЭУ), факультет социальных технологий (кафедра социологии); Российский университет дружбы народов (РУДН), факультет гуманитарных и социальных наук.

За годы преподавания курса социологии в высшей школе накоплен значительный опыт обучения студентов. Однако по-прежнему актуальны вопросы: с чего начинать социологическое образование в вузах и что особенно важно для обучения социологов, которое осуществляется и в региональных университетах, таких как Алтайский государственный университет, факультет социологии; Новосибирский государственный университет, кафедра общей социологии экономического факультета; Якутский государственный университет, кафедра социологии; Казанский финансово-экономический институт, кафедра философии и социологии; Нижегородский государственный университет, кафедра общей социологии и социальной работы; Иркутский государственный университет, кафедра социологии и социальных наук; Омский государственный университет, кафедра социологии, политологии и регионоведения; Удмуртский государственный университет, факультет социальных коммуникаций и др.?

Для достижения принципиально нового уровня развития социологической науки и социологического образования необходима такая система информационного обеспечения, которая позволяет получать максимально полное представление о современном потоке научной информации по социологии на разных языках и различных носителях. Важную роль в отборе, обработке, анализе, поиске и распространении информации по социологии играет Институт научной информации по общественным наукам (ИНИОН) РАН, основная цель которого заключается в информационном обеспечении научных исследований в области социальных и гуманитарных наук, в первую очередь – фундаментальных, а также в информационной поддержке учебного процесса в вузах. Информационная продукция ИНИОН в области социологии представлена как традиционными изданиями различного типа (научно-аналитическими обзорами, указателями литературы, реферативными сборниками), так и электронными информационны-

ми ресурсами, прежде всего, библиографическими и реферативными базами данных.

В основу формирования информационного массива для включения в БД различного типа положены такие принципы, как экспертный анализ и отбор наиболее значимых статей из журналов и сборников, полнота представления монографических изданий, многоаспектность содержательной обработки документов, отобранных в результате информационной фильтрации потока литературы по социальным и гуманитарным наукам, а также их видовое и языковое разнообразие. При этом отбор научных публикаций по социологии, глубина и многоаспектность отражения содержания социологической информации в отраслевой БД во многом зависят от целей и задач данной науки в условиях социально-экономических преобразований в стране, когда меняются решительно все стороны жизни – экономика, политика, социальные институты, система ценностей, быт. Социология призвана найти ответы на следующие вопросы: Каково происхождение общества, в котором мы живем? Что происходит в обществе в настоящее время? Куда мы движемся? И: что делать? Это все реалистичные вопросы, так как имеется реальный социальный мир с реальными качествами, населенный реальными людьми, которые коллективно произвели прошлое и чьи каузальные возможности уже сейчас задают контуры будущего (<http://isa-global-dialogue.net/wp-content/uploads/2013/07/v3i1-russian.pdf>).

В настоящее время, по мнению специалистов, перед социологами встают следующие задачи: во-первых, понять радикальные трансформации в обществе на основе применения различных теоретических и методологических подходов и, во-вторых, найти максимально адекватную связь между теорией и повседневной социальной практикой. Кроме того, для выхода на уровень общепринятых научных стандартов мировой социологии важным является освоение разностороннего языка социологической науки, позволяющего воспринять новые идеи, понятия, подходы и концепции. Сегодня, когда все большее значение приобретает практическая направленность профессии социолога, все более важным становится целостное видение социологических категорий и понятий, отражающих социальное развитие и функционирование человека в обществе.

ОБЩАЯ ХАРАКТЕРИСТИКА БАЗЫ ДАННЫХ ПО СОЦИОЛОГИИ

Формирование базы данных (БД), содержащей социологическую информацию, было начато в ИНИОН РАН в конце 1970-х гг. в процессе комплексной автоматизированной обработки литературы по общественным наукам, прежде всего, при подготовке и выпуске библиографических указателей литературы (текущих и ретроспективных). Реализация этих сложных и взаимосвязанных задач обусловила специфические, а в ряде случаев и уникальные особенности БД по социальным и гуманитарным наукам, наиболее важными из которых являются: во-первых, проблемно-тематическое соответствие входного информационного потока проблематике научных исследований в об-

ласти социальных и гуманитарных наук в целом и социологии, в частности; во-вторых, разработка комплекса лингвистических средств для многоаспектного описания предметного содержания самых разных видов изданий для последующего хранения, поиска и распространения социологической информации с помощью различных БД; в-третьих, широта языкового охвата документов, возможность получения информации на двух и более языках.

Поиск информации по социологии для информационной поддержки научных исследований и справочно-информационного обслуживания социологов различных организаций осуществляется в ИНИОН с начала 1980-х гг. Социологическая информация содержится в различных БД – отраслевых, сводных (интегрированных), проблемно-ориентированных, общий объем которых превышает 3,5 млн документов. Наиболее полно поток документов по социологии представлен в интегрированной БД «Философия и социология», включающей почти 435 тыс. документов. Массив социологической информации в этой БД превышает 140 тыс. документов. Ежегодный прирост информации по социологии составляет около пяти тысяч документов.

В настоящее время социологическая информация доступна потребителям информации на сайте ИНИОН в Интернете (www.inion.ru), а также на оптических дисках (DVD) на основе поисковой системы WebIRBIS™, предназначенной для многоцелевой обработки больших, в том числе полнотекстовых документов. Кроме того, для поиска научной информации по социологии могут использоваться традиционные информационные издания, прежде всего, библиографические указатели литературы, которые выпускаются в автоматизированном режиме. Так, с 1993 г. ИНИОН ежемесячно издает библиографический указатель «Новая литература по социальным и гуманитарным наукам. Философия. Социология», который является продолжением начатых еще в 1946 г. указателей «Новая советская литература по общественным наукам. Философские науки» и «Новая иностранная литература по общественным наукам. Философия и социология». Значительный вклад в создание отраслевой БД по социологии внесли такие известные библиографы, как О.С. Ли, С.И. Шапарина, Е.И. Серебряная, Б.П. Гинзбург, И.Л. Бельский, А.В. Борисов, А.К. Воскресенский и др.

При формировании информационного массива по социологии из общего документопотока по социальным и гуманитарным наукам отбираются документы по основным структурным разделам социологии, включая историю социологии, организацию современной научно-исследовательской работы, направления и школы современной социологии, методологию социологии, методику и технику социологических исследований, а также документы, описывающие общество как систему, социальные отношения и процессы, работы, посвященные социальным классам, обществам и группам. Наконец, учитываются документы по социологии конкретных сфер социальной жизни, социальных явлений и институтов: социология экономики, социология государства, права и политики,

социология международных отношений, войны и мира, социология общественного сознания, социология языка, социология морали, социология семьи и брака и т. д.

При этом в БД по социологии вводятся различные типы документов: монографии, сборники статей, отдельные статьи и рецензии из сборников и журналов, многотомные издания, учебные издания, материалы научных конференций, библиографические и справочные издания, авторефераты диссертаций. Важно отметить, что около 75 % информационного потока по социологии составляют статьи из научных журналов и сборников. Что касается научных работ, депонированных в ИНИОН, то с 1994 г. они вводятся в специализированную БД «Депонированные научные работы».

Анализ языкового состава информационного массива по социологии показывает, что социологическая информация в отраслевой БД представлена на более чем 40 языках, в том числе и на восточных языках. Почти 80 % данного массива составляют документы на русском и английском языках (50,2 и 29,1 % соответственно). Количество документов на немецком и французском языках – чуть более 5%. Среди других языков, на которых поступает информация по социологии в отраслевую БД, следует выделить болгарский, польский и китайский языки.

ЛИНГВИСТИЧЕСКИЕ СРЕДСТВА ПО СОЦИОЛОГИИ

Комплекс лингвистических средств по социальным и гуманитарным наукам обеспечивает формальную и содержательную обработку многоязычного потока документов и запросов с учетом специфики отдельных наук и научных дисциплин и, в первую очередь, с учетом особенностей понятийных и терминологических систем в различных языках и научных школах, направлениях, течениях. Для индексирования документов и запросов, а также для уточнения результатов поиска, расширения или сужения информационной области поиска используются, прежде всего, Рубрикатор Автоматизированной информационной системы по общественным наукам (АИСОН) [1] и Информационно-поисковый тезаурус ИНИОН по социологии [2].

Отраслевой раздел Рубрикатора АИСОН по социологии (А04) содержит более 130 рубрик и под рубрик. В его основу положена научная классификация отрасли, на основе которой определены тематический охват, структура и основные понятия философии, границы отрасли. В целом отраслевой Рубрикатор отражает современную структуру социологической науки и раскрывает проблемно-тематический профиль БД по социологии. Именно с помощью отраслевого Рубрикатора возможно быстрое выделение информационного массива по социологии из интегрированной БД «Философия и социология».

Проблемно-тематическая структура БД по социологии, а также распределение документов по рубрикам отраслевого Рубрикатора и типам изданий показаны в Таблице.

**Распределение документов в отраслевой базе данных по социологии
по рубрикам Рубрикатора ИНИОН (данные на 1.06.2015)**

Код рубрики	Наименование рубрики Рубрикатора ИНИОН	Кол-во документов		
		общее	статьи	монографии
A04	Социология	143995	94376	48726
A0401	Общие вопросы	11328	2573	8702
A0411	Общие проблемы современной социологии	2678	1997	616
A0415	Методология социологии. Методика и техника социологических исследований	4206	3283	879
A0421	Общество как система. Социальные отношения и процессы	24068	14898	9020
A0441	Социальные классы, общности и группы	23267	14741	7384
A0451	Социология сфер социальной жизни, социальных явлений и институтов	94327	62227	31607
A0461	Социология личности и поведения	8489	5793	2635
A0471	Историческая и региональная социология	7331	4400	2884
A0481	Прикладная социология	694	447	212
A0491	История социологии	2971	2324	582
A0494	Библиографические и справочные издания	1068	72	991

Каждая рубрика представлена в отраслевом Рубрикаторе как в словесной форме (на русском и английском языках), так и с помощью буквенно-цифровой нотации (кода), которая отражает иерархическое построение рубрикатора. Например, рубрика «Социальные классы. Социальная стратификация и мобильность» имеет код A044121, который фиксируется в библиографической записи документов, причем буква А вводится на латинице.

Иерархия рубрик Рубрикатора предстает в следующем виде:

- A04 Социология
- — A0441 Социальные классы, общности и группы
- — — A044121 Социальные классы. Социальная стратификация и мобильность.

Анализ статистических данных Таблицы показывает, что большая часть информационного массива относится к таким рубрикам, как A0451 «Социология сфер социальной жизни, социальных явлений и институтов» – почти 95 тыс. документов, A0421 «Общество как система. Социальные отношения и процессы» – свыше 24 тыс. документов, A0441 «Социальные классы, общности и группы» – более 23 тыс. книг и статей. Значительно меньше информации включено в отраслевую БД по таким рубрикам, как A0481 «Прикладная социология» – свыше 200 монографий и около 500 статей.

Другим важным лингвистическим средством для индексирования документов и запросов является информационно-поисковый тезаурус (ИПТ) по социологии, формирование которого осуществлялось в течение более 20 лет на основе анализа результатов содержательной обработки документов по социологии, а также в процессе изучения понятийного и терминологического состава отечественных и зару-

бежных социологических словарей и энциклопедий [3-5]. Тематический охват понятий в ИПТ базируется на представлении о структуре научного знания в области социологии, отраженной в отраслевом разделе Рубрикатора.

В настоящее время ИПТ по социологии, содержащий более 2 330 терминов, применяется, во-первых, в традиционном виде – как печатное издание, во-вторых, как лингвистическая база данных на оптическом диске или в Интернете, в-третьих, как структурированный словарь, встроенный в систему поиска информации по социологии в отраслевой БД, распространяемой на DVD. Тезаурус выполняет, прежде всего, функцию терминологического контроля при вводе документов и запросов в систему, обеспечивая единообразное представление в системе близких по содержанию документов. Понятия, необходимые для описания основного содержания документов, представлены в ИПТ дескрипторами - лексическими единицами отраслевой терминологии, нормализованными по принятым в АИСОН правилам.

Словарная статья каждого дескриптора представляет собой перечень лексических единиц ИПТ, семантически связанных с ним определёнными типами отношений. Необходимо подчеркнуть, что во многих отраслях социальных и гуманитарных наук, в том числе и в социологии, строгая дифференциация отношений между терминами (понятиями) вообще вряд ли возможна. В связи с этим для целей содержательной обработки документов и запросов было признано целесообразным лишь фиксировать в тезаурусе наличие иерархических (обозначение ВТ – выше и NT – ниже) и ассоциативных отношений (RT) между дескрипторами без уточнения их особенностей (например, род-вид, целое-часть, общее-частное и пр.). В качестве ассоциативных отношений между дескрипторами в ИПТ фиксируются, прежде всего, те отно-

шения, которые полезны как для раскрытия содержания понятия или определения его места в понятийной системе науки в целом, так и для расширения (или уточнения) границ поиска по информационным запросам или модификации поисковой стратегии.

Некоторые термины включены в ИПТ как условные синонимы (простые или сложные) с отсылкой к терминам, разрешенным для использования (дескрипторам). При индексировании документов или запросов синонимы заменяются на соответствующие дескрипторы. Например, термин «социальные страты» является условным синонимом, и при индексировании вместо него применяется дескриптор «социальные слои».

Приведем пример словарной статьи дескриптора «социальные слои» из ИПТ по социологии:

Дескриптор:	социальные слои
BT (вышестоящие термины):	общности
NT (нижестоящие термины):	социальные группы
	аристократия
	интеллигенция
	касты
	сословия
	средние слои
	элита
RT (ассоциативные термины):	социальная стратификация
	социальные классы
UF (используй вместо)	социальные страты
РУБ (рубрики):	A045143; A0461

В последние десятилетия опубликовано немало словарей по социологии, из них можно выделить, во-первых, «Большой толковый социологический словарь» [4], в котором особое внимание уделяется анализу понятийного аппарата социологии, методологическим и методическим требованиям по подготовке и реализации эмпирических социологических исследований, и, во-вторых, подготовленный специалистами социологического факультета РГГУ «Тезаурус социологии» [5], в котором терминологический массив упорядочен в соответствии с логикой взаимосвязей отдельных понятийных комплексов, образующих когнитивную структуру современного социологического знания. Нужно отметить, что именно этот тематический словарь-справочник может стать своего рода интеллектуальным навигатором не только для учебных программ по социологии, но и для обработки и поиска социологической информации.

Важной особенностью дескрипторного ИПЯ, используемого при содержательной обработке документов по социологии, является включение в поисковый образ документа так называемых аспектных дескрипторов, которые позволяют конкретизировать уровень (аспект) раскрытия основного содержания документа. В связи с тем, что один и тот же объект социологического исследования может рассматриваться в разных аспектах, использование аспектных дескрипторов имеет особое значение для повышения точности поиска. В АИСОН используются 8 аспектных дескрипторов, которые обозначаются одно сим-

вольными буквенными индексами. Так, аспектный дескриптор Т, обозначающий теоретический аспект рассмотрения проблемы (темы), позволяет быстро найти публикации теоретического характера. Быстрый поиск документов, в которых изучение того или иного объекта осуществляется на уровне страны обеспечивает аспектный дескриптор С, а на международном уровне – аспектный дескриптор М.

Аспектные дескрипторы являются средством повышения точности описания содержания документа и, соответственно, поиска в базах данных. Это особенно важно при больших объемах выдачи информации в результате поиска социологической информации. В таком случае аспектные дескрипторы обеспечивают конкретизацию в соответствии со своим содержанием. Так, если в библиографической записи документов содержатся дескрипторы «Россия» и «политические партии», то включение аспектного дескриптора «Л», обозначающего локальный аспект рассмотрения темы, позволяет при поиске из большого информационного массива отобрать лишь те документы, в которых речь идет о политических партиях только в российских регионах.

ТЕХНОЛОГИЯ СОДЕРЖАТЕЛЬНОЙ ОБРАБОТКИ ДОКУМЕНТОВ

Содержательная обработка документов по социологии осуществляется в два этапа. Первый этап – анализ документа с целью выявления его основного информационного содержания. Многоаспектное отражение основного содержания предполагает обязательный просмотр документа, при котором особое внимание уделяется не только заглавию, но и оглавлению, прикнижной аннотации, введению, заключению, резюме, предметному указателю в книге.

Второй этап содержательной обработки включает описание основного содержания документа с помощью комплекса лингвистических средств и формирование поискового образа документа с помощью дескрипторов, ключевых слов, аспектных дескрипторов, кодов рубрик. Обязательным элементом описания содержания документов на иностранных языках является аннотация. При ее составлении используется естественный язык, элементы которого могут применяться при формировании информационных запросов и уточнении результатов поиска. В системе предусмотрены разные типы аннотаций, наиболее распространенными из которых являются: 1) перевод заглавия документа; 2) расширение, уточнение или изменение заглавия при переводе на русский язык в случае его малой информативности на иностранном языке; 3) дополнение заглавия путём включения определённых формальных характеристик документа. Аннотация для документов на русском языке является факультативной и составляется в случае, когда название документа недостаточно информативно.

Каждый документ в БД представлен в форме библиографической записи, в структуре которой содержатся фрагменты формального и смыслового описания документа. Структура и форма представления библиографической записи соответствуют «Системе

стандартов по информации, библиотечному и издательскому делу" (СИБИД), согласованной со стандартами ИСО. Формат библиографической записи, используемой при обработке документов по социологии в АИСОИ, относится к полным универсальным форматам семейства MARC и обеспечивает многофункциональные входные и выходные формы библиографической записи. Общее количество полей, в которые вводятся данные о документе, превышает 200. Элементы некоторых полей объединены в так называемые инверсные файлы (словари, индексы), обеспечивающие проведение быстрого поиска.

Приведем пример библиографической записи статьи по социологии на английском языке.

Тип документа: статья Язык: английский Шифр: 15517642 Social structure and personality under conditions of radical social change: A comparative analysis of Poland and Ukraine / Koh M.L., Slomeczynski K.M., Janicka K. et al. // Amer. sociol. rev. - Wash., 1997. - Vol. 62, N 4. - P. 614-638. Bibliogr.: p. 638. Аннотация: Влияние социальной стратификации на развитие личности: сопоставительный анализ данных исследований в Польше и Украине. Ключевые слова: Польша; Poland; Украина; Ukraine; развитие личности; personality development; социальная стратификация; social stratification; сопоставительный анализ; comparative analysis Рубрики ИНИОН: A0461; A044121

Библиографические записи документов, созданные в процессе содержательной обработки документопотока по социологии, вводятся в так называемую буферную базу данных, на основе которой в автоматизированном режиме формируются БУ и вспомогательные указатели к ним. После подготовки указателя к изданию все библиографические записи документов включаются в поисковые БД (сводные, отраслевые, проблемно-ориентированные).

Важно отметить, что многоаспектность и политематичность документов по социологии, возможность рассмотрения одного и того же объекта исследования с точки зрения различных научных дисциплин приводят к условности отнесения целого ряда документов к той или иной отрасли. Некоторые документы, определённые как социологические, часто относятся и к другим отраслям, например, к философии, истории, науковедению или религиоведению. В связи с этим в процессе содержательной обработки документопотока часто происходит дублирование информационных материалов в нескольких отраслевых БД. Библиографические записи одного и того же документа, включаемого в разные базы данных, могут различаться с точки зрения содержательной обработки (прежде всего, аннотациями и дескрипторами), поскольку в каждой отрасли социальных и гуманитарных наук используются свои лингвистические средства. Анализ социологической информации в различных БД ИНИОН показывает, что общий объем документов по тем или иным аспектам социологии превышает 45 тыс. записей.

ПОИСК ИНФОРМАЦИИ ПО СОЦИОЛОГИИ В БАЗАХ ДАННЫХ ИНИОН

Поиск информации по социологии проводится, прежде всего, в отраслевой БД на основе «Руководства пользователя» [6] и «Инструкции по поиску», доступной на сайте ИНИОН. Для формирования запросов, поиска, просмотра, редактирования и выдачи информации по запросу используются различные возможности поисковой системы WinIRBIS. Проблемно-тематический поиск проводится по основным полям библиографических записей документов, для каждого из которых формируется свой словарь. При этом при каждом элементе словаря указывается количество документов, в описание которых он включен (код рубрики, термин, слово или словосочетание, дескриптор, фамилия автора и пр.). Поиск информации может проводиться как по отдельным полям библиографической записи, так и по их комбинации. При построении запросов (поисковых предписаний) используются знаки правостороннего и левостороннего усечения поисковых элементов, логические операторы ИЛИ, И, НЕ, а также круглые скобки.

В базе данных по социологии на оптических дисках (DVD) применяются встроенные в систему лингвистические средства (отраслевой рубрикатор и ИПТ), которые обеспечивают возможность расширения (или сужения) границ поиска, в том числе эвристического. Кроме того, в системе установлен двуязычный (русско-английский) интерфейс, который облегчает доступ к информации по социологии зарубежным пользователям.

Результативность поиска информации по социологии во многом зависит от выбора поисковых полей, стратегии и режима поиска, а также от использования различных лингвистических средств. Так, поиск трудов О.Конта проводится по полю авторов, а документов, в которых работы О.Конта анализируются, по полю «Общий словарь» или полю «Ключевые слова». Следует отметить, что в системе предусмотрена возможность совместного применения нескольких стратегий поиска, обеспечивающих итерационное повышение эффективности поиска путем генерации новых поисковых предписаний на основе последовательного редактирования (или комбинирования) результатов уже выполненных поисков.

Для повышения точности поиска при формировании запроса и поиске используются дескрипторы ИПТ по социологии, а также слова и словосочетания из релевантных документов, отобранных на предварительном этапе. Увеличение точности поиска может быть достигнуто за счет комбинированного поиска по кодам рубрик и дескрипторам, ограничения выдачи по аспекту рассмотрения темы, а также по формальным признакам — году издания, типам документов, языку документов. Например, по коду Рубрикатора A0494 можно быстро найти справочные и библиографические издания по социологии, среди которых по типу издания можно отобрать свыше 990 книжных изданий.

Одним из наиболее важных требований, предъявляемых потребителями к научной информации по

социологии, является ее полнота. Основная стратегия поиска социологической информации, которая ориентирована на информационную полноту, предусматривает проведение поиска по кодам рубрик отраслевого рубрикатора. При этом выделение базы данных по социологии из интегрированной БД «Философия и социология» производится путем поиска документов, относящихся к рубрикам отраслевого рубрикатора – по коду А04*. Возможна и другая стратегия, основанная на исключении из результатов поиска документов по философии, которые имеют в библиографической записи код рубрики А02* или А24*.

Один из способов увеличения полноты поиска социологической информации – это включение в запрос синонимичных или близких по значению терминов, содержащихся в релевантных документах, а также вышестоящих и/или ассоциативных дескрипторов, отобранных из встроенного в систему поиска ИПТ по социологии.

Система поиска социологической информации, реализованная в ИНИОН, позволяет потребителям самим запрашивать информацию из системы, пользуясь достаточно простым интерфейсом. Для профессионального поиска в системе предусмотрены три режима, из которых наиболее часто используется так называемый режим поиска «по образцу», обеспечивающий выбор поисковых терминов из словарей (тезаурусов) и логическую связь между элементами запроса. Взаимодействие пользователя с системой реализуется с помощью многооконного интерфейса, позволяющего формировать поисковые запросы любой сложности, редактировать результаты поиска, модифицировать поисковую стратегию, а также представлять найденную информацию в различных форматах (в том числе – транслитерированном).

Одним из наиболее сложных этапов информационного обеспечения потребителей социологической информации является анализ их информационных потребностей, определение основных требований к ретроспективной глубине и объему выдачи необходимой информации и, в особенности, выявление тех запросов, удовлетворение которых невозможно без новой информации и выбора релевантных информационных ресурсов, в которых ее можно отыскать. Для эффективного информационного поиска научных публикаций по социологии важным является применение различных стратегий, позволяющих обеспечивать как информационную полноту, так и точность поиска.

Необходимо подчеркнуть, специфика АИСОН заключается в том, что она ориентирована на различный уровень подготовки пользователей и позволяет строить информационный запрос разными способами с использованием всего комплекса лингвистических средств. Для повышения полноты поиска в системе предусмотрен механизм так называемого эвристического поиска, осуществляемого по запросу, который формируется автоматически на основе анализа релевантных документов, отобранных на первом этапе поиска.

Кроме того, в АИСОН предусмотрена возможность использования различных БД, в которых содержится социологическая информация, – отраслевых, сводных, проблемно-ориентированных, специализированных, реферативных. Такой поиск позволяет найти труды известных социологов, выделить работы продуктивных авторов, определить их научные интересы в разные периоды научной деятельности, проследить динамику развития их научных подходов, а также установить ключевые направления исследований в различных областях социологии, представляющие наибольший научный интерес для изучения. Следует отметить, что область поиска информации по социологии может быть расширена путем последовательного отбора документов по одному и тому же запросу из двух, трех или четырех БД. Так, при поиске социологической информации в базах данных по экономике и демографии, историческим наукам, государству и праву, науковедению только по ключу «социол*» можно найти более 15 тыс. документов.

Приведем примеры документов по социологии из различных БД ИНИОН.

Пример библиографической записи статьи на французском языке из БД по экономике и демографии:

Тип документа: статья	Язык: французский
Шифр: 04927642	
Davoine, L. L'economie du bonheur // Rev. econ. - P., 2009. - Vol. 60, N 4. - P. 905-926. Bibliogr.: p. 922-926. Рез. англ.	
Аннотация: Возможности использования данных социологических обследований о субъективном ощущении благосостояния как критерия социального развития общества и разработки государственной социальной политики. Страны ОЭСР.	
Ключевые слова: социальное развитие; social development; социальная политика; social policy; ОЭСР страны; OECD Countries	
Рубрики ИНИОН: А065855	

Пример библиографической записи монографии на немецком языке из БД по религиоведению:

Тип документа: монография	Язык: немецкий
Шифр: 069269212	
Zur Soziologie des Katholizismus / Hrsg. von Gabriel K., Kaufmann F.-X. – Mainz: Matthias-Grunewald-Verl, 1980. - 249 S. Bibliogr.: S. 226-240. Indices: S. 243-249.	
Аннотация: Социологические исследования проблем католической церкви в современном мире (культура, государство, социальная философия).	
Ключевые слова: социология религии; sociology of religion; католицизм; Catholicism; церковь и государство; church and state; религия и культура; religion and culture; социальная доктрина	
Рубрики ИНИОН: А21156121; А213155	

Пример библиографической записи статьи на немецком языке из БД по языкознанию:

Тип документа: статья **Язык:** немецкий
Шифр: 04477642
Kertesz, A. Chomsky, Wittgenstein, Bloor : Zum Problem einer wissenssoziologischen Metatheorie der Linguistik // Ztschr. für philos. Forschung. - Meisenheim/Glan, 1990. - Bd 44, H. 1. - S. 68-84
Аннотация: Дискуссия о социологических метатеориях языка.
Ключевые слова: философия языка; philosophy of language; социолингвистика; sociolinguistics
Рубрики ИНИОН: A16210755; A162127

Таким образом, информация по социологии представлена практически во всех отраслевых базах данных ИНИОН, однако ее доля в каждой из них различна. В связи с этим для обеспечения полноты поиска информации по социологии необходимо учитывать особенности всего информационного потока, поступающего в АИСОН, предметно-тематическая структура которого описывается Рубрикатором ИНИОН. При этом поиск может проводиться не только в отраслевых БД, но и в сводных базах данных (в частности, в БД текущих поступлений), обеспечивающих возможность ретроспективного поиска информации на заданную глубину (в частности, за последние 3 года). Например, при поиске в сводной БД за 2011-2013 гг. будет найден массив информации, содержащий свыше 3,5 тыс. документов.

Приведем пример библиографической записи монографии по социологии литературы из этой сводной БД:

Тип документа: монография **Язык:** английский
Шифр: 071961212
Sociological insights of great thinkers: Sociology through lit., philosophy, a. science / Edling Chr. a. Rydgren J. ed. - Oxford etc.: Praeger, 2011. - X, 370 p. Bibliogr.: p.323-344. Ind.: p.345-361.
Аннотация: Связь социологии литературы с философскими и научными исследованиями: сборник статей.
Ключевые слова: социология литературы; sociology of literature; философия; philosophy; научные исследования; scientific research
Рубрики ИНИОН: A045157

Поиск в сводных БД целесообразно проводить, прежде всего, по запросам политематического или межотраслевого характера, ориентированным, например, на высокую полноту поиска информации по теме исследования или поиск документов на двух и более языках.

Приведем пример библиографической записи сборника статей на английском и русском языках из сводной БД новых поступлений:

Тип документа: монография **Язык:** английский
русский **Шифр:** 026291312
Understanding the man - understanding the world / Eds. Kwiatkowska G.E., Bera R. - Lublin: Wydaw. Uniw. Marii Curie-Sklodowskiej, 2012. - 120 p.
Текст англ., рус. Библиогр. в конце ст.
Аннотация: Проблемы понимания мира через понимание человека: сборник статей.
Ключевые слова: понимающая социология; understanding sociology; понимание; comprehension; мировая система; world system; человек; man
Предм. рубрики: Общественные науки; Сборники; 2011–
Рубрики ИНИОН: A041121

Реализация политематического подхода к поиску информации по социологии возможна лишь на основе использования новых информационно-коммуникационных технологий и доступа к полнотекстовым электронным ресурсам (отечественным и зарубежным). Одним из средств повышения качества научно-информационного обеспечения потребителей информации по социологии является создание интегрированной базы данных реферативной информации, которая формируется в ИНИОН более десяти лет и включает полные тексты рефератов по многим отраслям социальных и гуманитарных наук. В настоящее время они доступны на сайте Национальной электронной библиотеки (www.elibrary.ru).

Как показывает опыт информационного обслуживания ИНИОН, пользователями, которые способны самостоятельно проводить качественный автоматизированный поиск социологической информации, являются лишь некоторые категории потребителей. Значительная часть тех, кому необходима научная информация по социологии, часто или не знают о существовании отраслевой БД в АИСОН, или по рекомендации, например, научных руководителей обращаются в информационную службу ИНИОН через Интернет (заказы на поиск информации оформляются на сайте ИНИОН), полагая, что именно информационные специалисты способны в полной мере удовлетворить их информационные потребности, используя все разнообразие информационных ресурсов, лингвистических средств и механизмов поиска.

ЗАКЛЮЧЕНИЕ

В настоящее время информационная поддержка и справочно-библиографическое обслуживание потребителей социологической информацией происходит в условиях значительного роста информационных ресурсов по социологии в электронной форме, появления новых информационно-коммуникационных технологий, широкого доступа пользователей различных категорий к ресурсам Интернета (отечественным и зарубежным). Важно отметить, что на базе интернет-технологий стала развиваться новая форма организации информационных ресурсов и предоставления услуг — отраслевые порталы и интернет-указатели социологических ресурсов. Сегодня проведение научных исследований, в том числе и в области социологии, невозможно без достаточно пол-

ного и оперативного информирования ученых о новейших достижениях науки в интересующих их областях. Представляется важным обеспечить доступ ученых, преподавателей и исследователей к отечественным и мировым информационным ресурсам не только с помощью баз данных ИНИОН.

Богатейшими информационными ресурсами по социологии обладает Институт социологии РАН (http://www.isras.ru/inform_resurs.html). Данные социологических исследований, проведенные сотрудниками Института за все годы его существования, собраны в Банке данных, в котором содержатся результаты более 700 исследований с 1966 г. Следует отметить, что Библиотека Института социологии РАН не только располагает уникальным фондом литературы по социологии, но и является единственным в России местом хранения полного комплекта отечественных изданий, начиная с 1969 г.

Важная информация по социологическим исследованиям с описанием методов исследования и их результатов содержится на сайте Центра социологических исследований МГУ им. М.В. Ломоносова (<http://www.opinio.msu.ru>). Результаты социологических исследований представлены также на сайте Центра социологических исследований Министерства высшего и профессионального образования РФ (http://www.informika.ru/windows/goscom/cinorgan/socio/first_pg.html).

При поиске информации по социологии могут быть полезны социологические ресурсы, ссылки на которые приводятся на сайте Научно-исследовательского социологического центра Нижегородского государственного университета (<http://nisots.nm.ru>).

Среди зарубежных источников социологической информации важно назвать, прежде всего, БД SocINDEX with Full Text, которая считается наиболее полной и высококачественной базой данных социологических исследований в мире. Она содержит более 2,4 млн документов, при создании библиографических записей которых был использован тезаурус по социологии объемом свыше 28,4 тыс. терминов. Важно отметить, что база данных включает аннотации к более 1500 основным журналам, начиная с изданий 1895 г., а также информацию из почти 420 профильных журналов и свыше 2,9 тыс. специально отобранных журналов. База данных SocINDEX with FullText предлагает полнотекстовые версии 826 журналов, начиная с изданий 1908 г. Специфическая особенность этой БД – наличие полнотекстовых версий книг и материалов конференций (<https://www.ebscohost.com/academic/socindex>).

Другим интересным источником социологической информации является ProQuest Sociology, предоставляющая доступ к рефератам документов (статей, книг, диссертаций, материалов конференций) из более 2 тыс. периодических изданий по социологии и социальной работе (<http://www.proquest.com/products-services/socioabs-set-c.html>).

Международная социологическая ассоциация (МАС), образованная под эгидой ЮНЕСКО в 1948 г., объединяет в настоящее время 60 национальных социологических ассоциаций (данные на 2012 г.), которые проводят научные исследования в рамках конвен-

циональных подходов и парадигм социологического знания. Несколько лет назад МАС приступила к реализации проекта Sociopedia.isa, основная цель которого состоит в создании онлайн-базы данных, охватывающей широкий спектр исследовательских тем, включая социальные конфликты, катастрофы, здоровье и болезни, диаспоры, проблемы памяти, мобильности, повседневности, транснациональных миграций, протестных движений и др. Важно отметить, что главная идея Sociopedia.isa заключается в новой концепции распространения социологических знаний за счет комбинирования возможностей Интернета для быстрой публикации обзорных статей и высокого качества подготовки научных электронных изданий, доступных членам МАС через сайт ассоциации или издательства SAGE (<http://www.isa-sociology.org/contr-sociopedia-isa.htm>).

СПИСОК ЛИТЕРАТУРЫ

1. Рубрикатор Автоматизированной информационной системы по общественным наукам (АИСОН) / РАН. ИНИОН. Фундам. Б-ка, Отд. науч.-библ. инфор., Отдел каталогизации и электрон. каталогов. Центр информатизации; сост. Г.С. Антонюк и др.; отв. ред. А.И. Слива, В.Л. Лейбович, М.Б. Шнайдерман. – М., 2012. – 284 с.
2. Информационно-поисковый тезаурус ИНИОН по социологии / РАН. ИНИОН. Центр информатизации. Фундаментальная библиотека / Б.П. Гинзбург при участии Р.Р. (Мдивани Классификационный указатель по ББК); ред.: Л.В. Гирко, О.С. Ли, Р.Р. Мдивани. – М., 2003. – 240 с.
3. Российская социологическая энциклопедия / под ред. академика РАН Г.В. Осипова. – М.: Норма-Инфра-М, 1998. – 672 с.
4. Большой толковый социологический словарь. В 2-х томах / пер. с англ. Н.Н. Марчук / Джери Д., Джери Дж. Т.1, А-О; Т. 2, П-Я. – М.: Вече, АСТ, 1999. – URL: <http://socioline.ru/pages/dzheri-dzheri-dzh-bolshoj-tolkovyj-sotsiologicheskij-slovar>
5. Тезаурус социологии. Книга 2. Методология и методы социологических исследований / под ред. Ж.Т.Тощенко. – М.: ЮНИТИ-ДАНА, 2013. – 416 с.
6. Библиографические базы данных. Руководство пользователя. Bibliographical Databases. User's Manual. – М.: ИНИОН РАН, 2001. – 34 с.

Материал поступил в редакцию 23.06.15.

Сведения об авторах

ШЕМБЕРКО Людмила Винцентовна – зав. сектором Института научной информации по общественным наукам (ИНИОН) РАН
e-mail: irichem@mail.ru

СЛИВА Алексей Иванович – кандидат исторических наук, зав. отделом ИНИОН РАН, Москва
e-mail: ais61@mail.ru

А.А. Ширяев

Управленческие информационные системы в научной сфере

Рассмотрены информационные системы, обеспечивающие организацию проведения научных исследований. Проведен количественный анализ систем, исследовано дублирование функционала, а также возможности по интеграции данных, собираемых различными порталами, с помощью проекта «Карта российской науки» с целью получения сведений обо всей научной деятельности в стране в разрезах, используемых Минобрнауки России при принятии решений. Такой подход позволит преодолеть фрагментированность собираемой информации об объекте и снизить бюрократическую нагрузку на ученого при заполнении форм заявочной и отчетной документации.

Ключевые слова: информационные системы, Карта российской науки, центры коллективного пользования, уникальные научные установки, Реестр РНТД

ВВЕДЕНИЕ

В России выделяются существенные средства на повышение эффективности и результативности научной деятельности. В частности, за 2010-2012 гг. внутренние затраты на исследования и разработки имели тенденцию к росту – с 610,4 млрд руб. до 699,9 млрд руб. соответственно (увеличение на 14,6%) [1, с. 17]. Однако при этом удельный вес затрат в ВВП снижается: с 1,27% в 2002 г., до 1,12% в 2012 г.¹ Необходимо отметить, что не выполняются плановые значения показателей Стратегии развития науки и инноваций в Российской Федерации на период до 2015 года, где указывалось, что к 2015 г. удельный вес внутренних затрат на исследования и разработки в ВВП составит 2,5%. В Указе Президента Российской Федерации от 07.05.2012 № 599 «О мерах по реализации государственной политики в области образования и науки» приведено другое целевое значение: 1,77% ВВП к 2015. К настоящему времени оба плановых значения не достигаются. Результативность научной деятельности, несмотря на рост финансирования, также не является существенной. По данным НИУ ВШЭ [1, с. 263-264] удельный вес российских публикаций в журналах, индексируемых в базах данных Web of science и Scopus, в 2012 г. составил 1,9% (в 2006 г. – 2,24%) и 1,74% (в 2006 г. – 1,97%) соответственно.

С целью повышения эффективности научной деятельности некоторые исследователи указывают на необходимость реформирования российской науки по иностранному варианту [2], увеличения финансирования науки в целом и РАН – в частности [3], а

также изучения противоречивости научно-технической политики [3]. Кроме того, имеется разработка, указывающая, что снижение результативности науки является следствием общего снижения интеллектуального потенциала общества [4].

На практике для развития науки в России выбран путь поддержки федеральных и национальных исследовательских университетов, создания национальных исследовательских центров и др. Предполагается, что укрупнение организаций повысит эффективность и результативность российской науки в целом, поможет в решении вопросов финансирования и определения приоритетов развития.

Другие исследователи анализируют вопрос бюрократии в научной сфере [5], либо обращают внимание на организацию процесса финансирования науки – переход от государственного задания на конкурсную модель финансирования [6]. Таким образом, критически оценивается научно-техническая политика в целом.

В настоящей статье мы рассматриваем обеспечивающие организацию научной деятельности информационные системы (далее – ИС): комплекующие информацию о деятельности научных организаций и через которые подаются заявки на получение финансирования проектов и собирается отчетная документация. Исследованию подлежит организующая составляющая ИС, что может выражаться в сборе информации о ресурсном потенциале и результативности научной деятельности в научном секторе в целом и в разрезах (для информирования лиц, принимающих решения), в снижении временных затрат ученых на заполнение документации и облегчении этого процесса, а также в степени доступности данных.

Проведенный анализ уровня бюрократии в науке (на примере федеральных целевых программ (далее –

¹ Данные Аналитического центра при Правительстве Российской Федерации. URL: <http://ac.gov.ru/files/attachment/4879.pdf>

ФЦП), грантов и государственного задания в сфере науки) показал недостаточную развитость информационных систем, обеспечивающих функционирование научного сектора. Было выявлено, что они применяют устаревшую технологию подготовки материалов в программе MsWord с последующим размещением в системе, не реализуют в полной мере возможности автоматической проверки вводимых данных, не используют заранее заданные списки, облегчающие заполнение заявочных и отчетных форм. Количество самих информационных систем, в которые необходимо вводить данные, значительно.

ИНФОРМАЦИОННЫЕ СИСТЕМЫ В НАУЧНОЙ СФЕРЕ

Условно можно выделить две группы информационных систем, обеспечивающих проведение научных исследований:

1) обеспечивающие подачу заявок на получение финансирования и принимающие отчетную документацию;

2) собирающие информацию о результативности научной деятельности и имеющемся научном потенциале.

Первую группу представляют:

- портал ФЦП «Исследования и разработки по приоритетным направлениям развития науки, технологий и техники» на 2014-2020 годы (далее – ФЦП ИиР) (<http://fcpir.ru/>);
- информационная система для подачи отчетной документации по ФЦП ИиР (<https://sstp.ru/>);
- сайты научных фондов: РГНФ, РФФИ, РНФ и грантов Президента Российской Федерации;
- информационная система формирования государственных заданий высшим учебным заведениям и научным организациям в сфере научной деятельности (<http://xn--80aahfgik3be4a.xn--p1ai/>).

Информационные системы второй группы представлены в таблице. Перечень не является исчерпывающим, поскольку Федеральное агентство научных организаций создает новые ИС, собирая данные о деятельности учреждений РАН.

Информационные системы, обеспечивающие сбор данных о научной деятельности

№ п/п	Название информационной системы	Электронный адрес
1	Единый реестр результатов научно-исследовательских, опытно-конструкторских и технологических работ гражданского назначения, выполняемых за счёт средств федерального бюджета	http://www.rntd.citis.ru/php/rntd4.php
2	Единая государственная информационная система учета результатов научно-исследовательских, опытно-конструкторских и технологических работ гражданского назначения (далее – ИС РНТД)	http://www.rosrid.ru
3	Государственной учет результатов научно-исследовательских, опытно-конструкторских и технологических работ гражданского назначения	http://intelpro.extech.ru/
4	Информационно-аналитические материалы по результатам анализа показателей эффективности образовательных организаций высшего образования*	http://miccedu.ru/monitoring/
5	Электронный каталог высокотехнологичного оборудования и объектов научного потенциала Российской Федерации	http://www.xn----7sbam6aiqfmx.xn--p1ai/
6	Современная исследовательская инфраструктура Российской Федерации: открытость, доступность, инновационность	http://www.ckp-rf.ru/
7	Карта российской науки	https://mapofscience.ru/
8	Развитие инновационной инфраструктуры в российских вузах	http://rii-vuz.extech.ru/
9	Учет и мониторинг малых инновационных предприятий научно-образовательной сферы	https://mip.extech.ru/
10	Федеральный реестр экспертов научно-технической сферы Минобрнауки России	https://reestr.extech.ru/experty/index.php
11	Проект повышения конкурентоспособности ведущих российских университетов среди ведущих мировых научно-образовательных центров	http://5top100.ru/
12	Федеральная система мониторинга результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы	http://www.sciencemon.ru/

* Информационная система содержит данные не только о показателях научной деятельности, но и об образовательной и прочей деятельности вузов.

Источник: составлено автором.

Размещение такого перечня на сайте Минобрнауки России или РАН повысит осведомленность научного сообщества об имеющихся возможностях, предоставляемых системами.

УЧЕТ РЕЗУЛЬТАТОВ НИОКР

Для выявления дублирования и противоречий рассмотрим следующие информационные системы.

1. Единый реестр результатов научно-исследовательских, опытно-конструкторских и технологических работ гражданского назначения, выполняемых за счёт средств федерального бюджета (<http://www.rntd.citis.ru/php/rntd4.php>).

2. Единую государственную информационную систему учета результатов научно-исследовательских, опытно-конструкторских и технологических работ гражданского назначения (<http://www.rosrid.ru>).

3. Государственный учет результатов научно-исследовательских, опытно-конструкторских и технологических работ гражданского назначения (<http://intelpro.extech.ru/>).

Информационные системы №1 и №3 имеют сходные названия. Однако следует отметить, что ИС №3 перестала работать с введением ИС №2. Целесообразность замены одной системы на другую, расходования дополнительных финансовых средств на ее создание при имеющейся аналогичной системе органами государственной власти не обосновывается.

Доступ к информационной системе №2 закрыт, хотя работы выполняются за счет средств бюджета, и общество имеет право на предоставление информации о направлениях расходования бюджетных средств и полученных за их счет результатах. Не вызывает сомнений, что эти результаты могут содержать эффект новизны, коммерческую и иную тайну. Однако совершенно неясна целесообразность проведения за счет средств федерального бюджета НИОКР с последующим полным закрытием всей информации от общества, экспертного, бизнес и научного сообществ. Доступ к отчетной документации по проектам имеют только представители органов государственной власти, проверяющие и контролирующие организации, а также исследователи, проживающие в московском регионе, поскольку они имеют возможность посетить Центр информационных технологий и систем органов исполнительной власти (далее – ЦИТиС), обеспечивающий как функционирование ИС №2, так и сбор всех отчетов о НИОКР по федеральным целевым программам. ЦИТиС находится, соответственно, в Москве.

Относительно информационной системы №1 неясно, как она соотносится с ИС №2. По всей видимости, информационная система №2 содержит больше подробной информации о результатах, полученных в рамках НИОКР. Из этой информации часть данных попадает ИС №1. К сожалению, потенциальные заинтересованные результатами НИОКР пользователи со стороны бизнеса могут воспользоваться информацией о патентах, полученных в результате выполнения работ, только если будут использовать ИС №1, но такой информацией они обладают и без этой системы.

Таким образом, ни научное, ни бизнес-сообщество не могут воспользоваться результатами НИОКР, выполненными за счет федерального бюджета, что ведет к невозможности внедрения их результатов в производство. Целесообразным была бы публикация если не полного отчета о НИОКР в свободном доступе, то хотя бы резюме проектов. Либо государство могло бы заказывать выполнение НИОКР по определённой теме, интересной бизнесу, за счет федерального бюджета с последующей безвозмездной передачей результатов для коммерциализации.

ИНФОРМАЦИОННЫЕ СИСТЕМЫ, ПРЕДОСТАВЛЯЮЩИЕ ИНФОРМАЦИЮ ОБ ИМЕЮЩЕМСЯ НАУЧНОМ ОБОРУДОВАНИИ

На данный момент функционируют две системы, предоставляющие информацию о научном оборудовании.

1. Электронный каталог высокотехнологичного оборудования и объектов научного потенциала Российской Федерации (<http://www.xn----7sbam6aiqfmx.xn--p1ai/>) (далее – ИС оборудования).

2. Современная исследовательская инфраструктура Российской Федерации: открытость, доступность, инновационность (<http://www.ckp-rf.ru/>) (далее – ИС ЦКП).

ИС оборудования, несмотря на название, определяющее ее как систему, предоставляющую информацию обо всем высокотехнологичном оборудовании в Российской Федерации, позиционируется как система об оборудовании, имеющемся в вузах и компаниях, в то время как научные организации на этом сайте не представлены.

ИС ЦКП содержит информацию только по центрам коллективного пользования научным и уникальным оборудованием (далее – ЦКП). Изначально в ней содержались сведения только об оборудовании центров коллективного пользования и уникальных научных установках, поддержанных Минобрнауки России по мероприятиям 1.8 и 5.2 ФЦП «Исследования и разработки по приоритетным направлениям развития науки, технологий и техники» на 2007-2013 гг. Однако позднее перечень организаций был расширен. Поскольку отдельные вузы имеют в своем составе также и центры коллективного пользования научным оборудованием, то сведения об их оборудовании включено в обе информационные системы, что ведет к дублированию функционала двух порталов.

При этом ни в одной из этих систем не реализована задача инвентаризации всего парка оборудования (научного, технологического, иного), имеющегося в вузах, НИИ и компаниях, выполняющих НИОКР, что не позволяет оценивать степень износа не просто машин и оборудования в научном секторе (такой показатель используется в статистике), а именно научного (измерительного и аналитического) и технологического оборудования.

Целесообразнее было бы создание единого портала, собирающего такую информацию и позволяюще-

го проводить аналитические исследования в разных разрезах: по областям науки, по классификатору обособления, по виду организации.

ИНФОРМАЦИОННЫЕ СИСТЕМЫ ФОНДОВ И ФЕДЕРАЛЬНЫХ ЦЕЛЕВЫХ ПРОГРАММ

Помимо информационных систем, собирающих информацию о научном потенциале и результативности научного сектора, существуют системы, целью которых является обеспечение подачи заявок и получения отчетности по грантам и конкурсам.

Необходимо отметить, что информационные системы государственных научных фондов имеют схожие цели и задачи – сбор заявок и получение отчетности. Однако в них отсутствует унификация пользовательского интерфейса. Они имеют свои формы, правила их заполнения, перечень необходимых подтверждающих документов. Это затрудняет для пользователя ориентирование в требованиях, предъявляемых системами.

Система сбора заявочной и отчетной документации по ФЦП ИиР является наиболее сложной, поскольку требует введения значительного количества данных по всем сотрудникам организации, принимающим участие в проекте, на который подается заявка: публикационная и патентная активность, индексы цитирования по каждой статье, представленной в заявке и др.

Таким образом, все указанные системы в заявках и отчетности так или иначе собирают идентичную информацию о публикационной активности исследователей и результативности проведенных работ в виде публикаций, патентов и др. Целесообразной в этом случае представляется унификация форм и процедур их заполнения или создание единого портала, на котором бы такая информация размещалась с возможностью ее дальнейшего использования в информационных системах фондов и ФЦП ИиР. Это могло бы значительно упростить заполнение заявок и отчетности учеными, поскольку требовало бы вместо заполнения всех форм для каждой отдельной ИС, указывать только номер учетной записи с последующей выгрузкой данных из единого портала данных о научной деятельности. Таким порталом мог бы стать проект «Карта российской науки» (далее – Карта).

КАРТА РОССИЙСКОЙ НАУКИ

Разработка Карты российской науки изначально задумывалась как создание портала, содержащего наиболее полную информацию о научной деятельности, материально-техническом обеспечении исследований, проводимых организациями, а также отдельными сотрудниками. Система позволяет собирать данные о публикационной активности.

Однако по состоянию на 26.02.2015 система все еще находится в стадии опытной эксплуатации. Сведения, содержащиеся в базе данных портала, актуальны на 2013 г. Информация о публикациях представлена сведениями из баз данных Web of Science и РИНЦ. База данных Scopus отсутствует на портале, хотя публикации, индексируемые в ней, оцениваются при подаче заявок на гранты и на участие в конкурсах. Информация из базы данных РИНЦ не актуализируется.

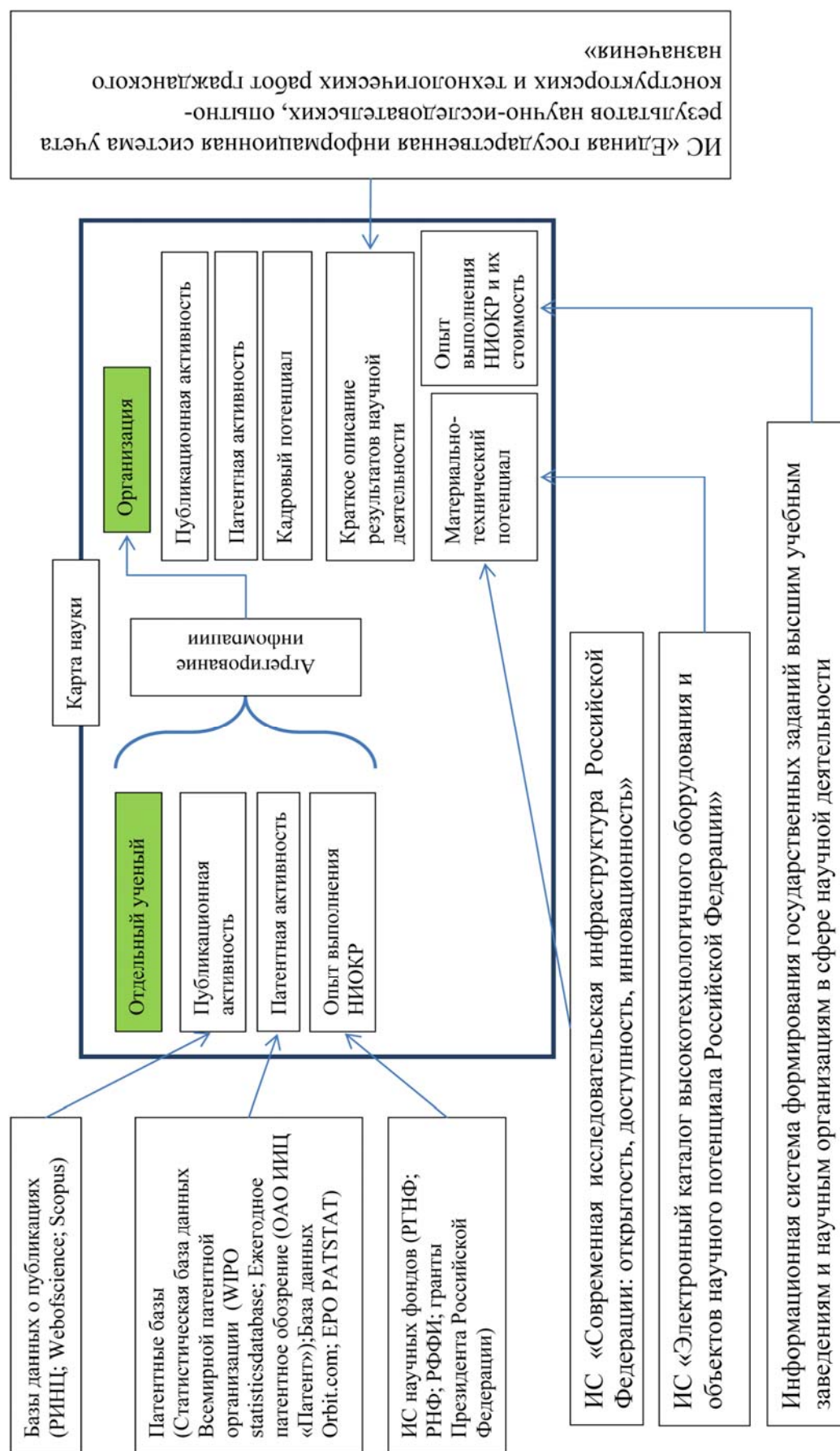
Фактически, Карта российской науки способствовала бы сокращению объемов заявочной и отчетной документации (по научным фондам, ФЦП ИиР), позволяя при ее заполнении указывать только идентификационный номер организации-заявителя и участников проекта, а прочие данные подгружались бы в заявку автоматически с портала Карты. Это могло бы снизить объемы документации, облегчить проверку заявочной и отчетной документации, снизить расходы на обеспечение проведения исследований. Такой механизм фактически не удалось реализовать в силу большого количества нареканий в некорректности данных, содержащихся в Карте², что приводит к невозможности их использования.

Помимо этого, Карта могла бы предоставлять информацию для обоснования принятия решений. Например, при выработке пороговых значений показателей публикационной активности молодых ученых в возрасте до 35 лет для допуска их заявок к участию в конкурсе научных фондов. Кроме того, Карта фактически является переписью российских ученых, а также инвентаризацией научного потенциала. Поэтому, в случае реализации возможностей Карты по сбору актуальной и более полной информации о научной деятельности, лица, принимающие решения, получили бы источник оперативной информации о научном потенциале и результативности научной деятельности в различных разрезах и с детализацией, недоступной официальной статистике. Аккумулируя информацию из многих источников, можно было бы получать данные о приборном парке в целом по стране, а также отдельно в вузах, НИИ, РАН, в ЦКП. Также были бы доступны сведения об источниках финансирования: ФЦП, государственное задание и фонды. Такие срезы информации используются в Минобрнауки России при принятии решений, однако данные приходится собирать отдельно под каждую конкретную задачу, поскольку официальная статистика их не предоставляет. Карта была бы более дешевой альтернативой официальной статистике и позволила бы получать более точные, чем официальные, статистические данные об объекте в нужных для принятия решений разрезах. Тем не менее, несмотря на большой потенциал проекта, он так и не был доведен до завершения.

На рисунке показана возможная схема работы проекта «Карта российской науки». Аккумулируя данные различных информационных систем, Карта способна стать порталом, содержащим информацию обо всей науке в России, и предоставлять:

- информацию о результативности и эффективности научной деятельности (таким образом исключается необходимость функционирования Федеральной системы мониторинга результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы);

² ОС ОНР: Карта российской науки (КРН) глазами ученых. Электронный доступ: <http://www.saveras.ru/archives/tag/%D0%BA%D0%B0%D1%80%D1%82%D0%B0-%D1%80%D0%BE%D1%81%D1%81%D0%B8%D0%B9%D1%81%D0%BA%D0%BE%D0%B9-%D0%BD%D0%B0%D1%83%D0%BA%D0%B8>



Предлагаемая схема агрегирования данных проектом «Карта российской науки»

Источник: составлено автором

- перечень наиболее публикуемых и цитируемых российских авторов (для отбора экспертов, которые могут участвовать в деятельности научных фондов и оценке результатов НИОКР, финансируемых государством);
- обоснование пороговых значений критериев оценки заявок, участвующих в конкурсах фондов и ФЦП;
- автоматическое формирование заявок на участие в конкурсах фондов и ФЦП, что снижает бюрократическую нагрузку;
- возможность получения оперативной статистической информации о научном секторе в более подробных разрезах, чем дает официальная статистика Росстата;
- краткое описание результатов НИОКР, полученных организацией;
- возможность прогнозирования основных показателей научно-технологического развития, в том числе на основе анализа исследовательских фронтов.

ЗАКЛЮЧЕНИЕ

Проведенный анализ российских управленческих информационных систем в научной сфере показал, что призванные обеспечивать организацию проведения научных исследований, облегчать работу ученых, в том числе по заполнению различных документов, а также собирать необходимую непротиворечивую информацию о российской науке, на практике эти системы не всегда выполняют поставленные задачи. Так, выявлены информационные системы, функционал которых может дублироваться. Несмотря на проведение научных исследований за бюджетные деньги, их результаты не являются прозрачными для общества, что выражается в закрытии информации, необходимости проходить авторизацию для получения доступа к отдельным сведениям о результатах проведенных работ.

Сбор данных о потенциале, результативности и эффективности деятельности российской науки в разрезах, фигурирующих в различных государственных документах³, которые используются для принятия решений, также затруднен. Сегодня под каждую задачу, являющуюся важной для министерств, создается отдельная информационная система. Такой подход приводит к невозможности собрать данные в целом об объекте, например, о научном оборудовании только центров коллективного пользования, а не обо всем научном и аналитическом оборудовании в стране. Получается, что информация, поступающая лицам, принимающим решения, мозаична. Карта науки

могла бы выступить интегратором данных, где бы аккумулировалась информация об оборудовании, кадрах, результативности деятельности в различных разрезах. Проект «Карта российской науки» мог бы стать связующим звеном между разрозненными порталами, обеспечив унификацию вводимых данных в заявочную документацию на получение финансирования и при заполнении отчетности, что избавило бы ученых от необходимости в разных системах вводить идентичную информацию, например при подаче заявок на гранты фондов и для участия в ФЦП ИиР.

Таким образом, система управления российской наукой, проанализированная через информационные системы о деятельности научного сектора, представляется недейственной и фрагментированной. Эффективность непосредственно научной деятельности снижается.

СПИСОК ЛИТЕРАТУРЫ

1. Индикаторы науки, 2014: статистический сборник. – М.: Национальный исследовательский университет «Высшая школа экономики», 2014. – 400 с.
2. Гохберг Л.М., Заиченко С.А., Китова Г.А., Кузнецова Т.Е. Научная политика: глобальный контекст и российская практика. – М.: Изд-во «Дом Высшей школы экономики», 2011. – 308 с.
3. Научная и инновационная политика. Россия и Мир. 2011-2012 / под ред. Н.И. Ивановой, В.В. Иванова. – М.: Наука, 2013. – 480 с.
4. Тодосийчук А.В. Интеллектуальный потенциал общества, результативность науки и экономический рост // Инновации. – 2010. – №1. – С. 35-42.
5. Резник Ю.М. Бюрократия и сообщества в науке: взаимодействие или противодействие? // Личность. Культура. Общество. Вып. 2. – 2013. – Том XV, № 78. – С. 8-22.
6. Гусев А.Б. Судьба федеральных целевых программ в сфере науки (проблемы системности и перспективы программно-целевого метода) // Общество и экономика. – 2014. – № 9. – С. 81-105.

Материал поступил в редакцию 30.04.15.

Сведения об авторе

ШИРЯЕВ Алексей Алексеевич – кандидат экономических наук, советник Аналитического центра при Правительстве Российской Федерации, Москва
e-mail: alexeishiraev@mail.ru

³ Например, о критических технологиях; имеющемся в организациях научном оборудовании, его возрасте.