

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 1. ОРГАНИЗАЦИЯ И МЕТОДИКА
ИНФОРМАЦИОННОЙ РАБОТЫ

ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 1

Москва 2014

ОБЩИЙ РАЗДЕЛ

УДК 316.776 - 043.87 : 655

А. В. Мельников, Э. П. Семенюк

Информационная революция и современная полиграфия

Прогресс человечества уже давно невозможен без качественного развития полиграфического производства: книга, как известно, является существенным атрибутом мировой цивилизации. За последние десятилетия важнейшей особенностью научно-технической революции стал ее информационный аспект, и он многогранно влияет на изменение природы современной полиграфии. Новейшие информационные технологии заметно пополнили арсенал традиционного книгопечатания и составляют основу использования нынешнего полиграфического оборудования.

Ключевые слова: научно-техническая революция, микроэлектроника, информационная революция, книга, книгопечатание, полиграфия, информационные технологии

Стремительный бег времени неизбежно рождает новые проблемы и задачи человечества. Сегодня в центре внимания общества мировой финансово-экономический кризис, борьба с голодом и бедностью, узел насущных ресурсных трудностей и экологических бедствий, череда кровопролитных вооруженных конфликтов на этнической, религиозной либо социальной почве, преодоление культа потребительства и господства бездуховности... Перечень этот, видимо, можно продолжить, но смысл сказан-

ного уже очевиден: еще в недавнем прошлом (скажем, на рубеже тысячелетий) многие из этих тенденций не были обозначены столь явственно, как сейчас. И завтрашний день, несомненно, принесет что-то новое.

Но вместе с тем есть проблемы большего масштаба, сопровождающие определенную социокультурную эпоху в жизни человечества. К ним, например, несомненно принадлежит тот факт, что уже давно мир живет в условиях научно-технической револю-

ции (НТР). Автором этого понятия, как известно, является английский физик, историк и социолог науки Джон Бернал (впервые термин был употреблён в его статье «Социальная функция науки» в 1938 г.). Как видим, с тех пор прошло уже 75 лет, и за это время понятие НТР приобрело широчайшую популярность в мировой науке и культуре. Это уже давно не просто удачная, звучная метафора, а достаточно строгое и емкое понятие современной философии науки и техники, отличающееся вполне определенным семантическим наполнением [1–3]. Более того, смысловое содержание этого концепта в значительной мере характеризует и важные черты всей планетарной цивилизации нашей эпохи.

Существо данного вопроса в том, что понятие НТР органично соотносится не только с областями науки, техники, технологий и материального производства (к которым имеет самое непосредственное отношение по своему содержанию и объему), но также и с широчайшей сферой духовной жизни общества. Иными словами, оно стало релевантным применительно к функционированию и развитию социума в целом. Именно поэтому самые различные по своей семантике явления, процессы и стороны в жизни человечества (включая и те, с которых мы начали эту статью) так или иначе должны осмысливаться в общем контексте научно-технической революции как одного из базисных детерминирующих факторов целой эпохи, ее сущности, характера, лица.

В данном случае нашей основной целью является анализ влияния НТР на прогресс полиграфии, давно уже представляющей собой специфичный и очень заметный по своей роли инструмент социокультурного развития. Этот круг вопросов, безусловно, связан с методологическими проблемами информатики: все, что имеет отношение к книге, к «галактике Гутенберга», вне сомнений, прямо и непосредственно характеризует глобальное информационное пространство. НТР в этом смысле означала гигантский рывок в будущее.

СОЦИАЛЬНАЯ РОЛЬ ИНФОРМАЦИОННОЙ ПЛОСКОСТИ НАУЧНО-ТЕХНИЧЕСКОЙ РЕВОЛЮЦИИ

Стратегические направления НТР анализируются исследователями достаточно давно. Конечно, динамика развития науки, техники, всего общества постоянно вносит в этот процесс (а также в его результаты) существенные коррективы, но при этом в итоге с необходимостью кристаллизуются определенные логико-семантические константы, своего рода реперные точки и линии анализа. Давно уже характерным стало методологическое соотнесение сущности НТР с особенностями современной науки, ее отличительными чертами и механизмами [4, 5]. И важнейший аспект этой проблематики имманентно связан с информационным подходом к действительности.

Следует подчеркнуть, что так было с самого начала НТР. Как известно, к числу принципиально важных научных ее истоков принадлежит становление кибернетики, в которой понятие информации играет роль одной из основных категорий. Теория информа-

ции очень скоро стала методологически необходимой составляющей всего могучего кибернетического комплекса научного знания. Соответственно этому в картине мира, формирующейся в эпоху НТР, информация начинает занимать одно из важнейших мест — наряду с материей и энергией [6, с. 201; 7, с. 136–137], и именно это положение легло в основу информационного подхода как существенного и даже необходимого компонента современной общенаучной методологии.

Но это, так сказать, сугубо научный (точнее, научно-методологический) срез проблемы. Гораздо более широкое, общесоциальное ее содержание стало понятным несколько позже, в 70-х гг. прошлого века, с началом микропроцессорной (компьютерной), или же информационной революции [8]. Этот глубоко своеобразный феномен, будучи неотъемлемой составляющей и важнейшим аспектом НТР, вместе с тем знаменовал собой развертывание ее особого, новейшего этапа — технологического. В свое время одному из авторов данной статьи уже довелось на страницах «НТИ» особо отмечать специфику этого этапа научно-технической революции и тот факт, что он теснейшим образом связан с информатикой [9]. Все началось с создания в 1971 г. микропроцессора — универсального управляющего устройства на микромином кристалле кремния.

Изобретение на основе микропроцессора больших и сверхбольших интегральных схем позволило заметно интенсифицировать развитие электронной вычислительной техники и одно за другим вызвало к жизни новые поколения компьютеров. Но сущность микропроцессорной революции отнюдь не только в прогрессе компьютерной техники как таковой. Широчайшее внедрение интегральных схем в основное технологическое оборудование промышленности и строительства, транспорта и связи, лесного и сельского хозяйства, в медицину и здравоохранение, полиграфию, образование, культуру, сферу обслуживания и т. п. открыло дорогу новейшим информационным технологиям практически во всех областях жизни общества. Эти преобразования без преувеличения стали информационной революцией планетарного масштаба.

Информационная плоскость научно-технической революции органично связана со всеми другими ее сторонами и направлениями уже потому, что жизнь и деятельность человека и социума попросту немислимы без информации. В наше время эта истина стала остро ощутимой буквально на каждом шагу.

Сущность НТР, как известно, заключается прежде всего в небывалом до этого слиянии и взаимопроникновении революционных преобразований одновременно в трех кардинальных сферах жизни общества — в науке, технике и производстве, составляющих сегодня основу всей социальной практики. В прошлом эти области (и соответственно изменения в них) были ощутимо разграничены, эпоха же НТР детерминировала их теснейшее взаимодействие и взаимовлияние. Именно на этой базе объективно сложилась и постоянно поддерживается в своем единстве органически целостная система «наука — техника — производство — социальная практика (в

полном ее объеме)», каждая из подсистем которой непрерывно влияет на все остальные революционизирующим образом [9, с. 2].

Информация буквально пронизывает каждую из этих сфер жизни общества. К тому же важнейшую роль играют разнообразные информационные потоки, постоянно связывающие все эти области между собой. Понятно, почему в культуре нашего времени получила широкое распространение метафора, когда информационную подсистему глобального социального организма называют аналогом нервной системы человека или животного.

Следует подчеркнуть, что осознание и познание информационного аспекта реальности началось значительно позже, чем изучение вещественно-энергетических факторов природы и жизни общества. История науки и техники последних столетий убеждает в том, что человечество постепенно накапливало знания для качественного скачка в понимании сущности информации и ее роли в мироздании — скачка, который пришелся именно на период НТР. Параллельно с этим исследование свойств электричества, а затем формирование электроники шаг за шагом открывали путь к созданию электронной вычислительной техники (необходимо также помнить о вкладе математической лингвистики в разработку алгоритмических языков программирования). Безусловно, это лишь важнейшие, наиболее очевидные штрихи долгого и сложного пути, но и они помогают осознать, что к результатам компьютерной революции вело взаимодействие многих отраслей научного знания. А другой, не менее существенной плоскостью информационного прорыва в науке стало формирование *философии информации* как особого направления прогресса философского знания [10, 11]. Наука и философия, как известно, при всех чертах своего духовного родства — генетического и содержательного — являются все же разными формами общественного сознания. И потому философско-методологическое осмысление столь сложного и онтологически многообразного феномена, как информация, открыло путь к более глубокому осознанию социальной роли информационного аспекта НТР.

Со временем в поле зрения исследователей попали и такие специфические (а вместе с тем — крайне важные социально) процессы, как компьютеризация, электронизация, медиатизация, наконец, информатизация общества [8, 12]. Все они производны от важнейших направлений научно-технической революции, развития и углубления ее сущности. С другой же стороны, эти процессы теснейшим образом связаны (и к тому же — многогранно) с прогрессом информатики и усилением ее влияния на современный мир [13].

Если в начале НТР (в 40-х гг. XX в.) абсолютным научным лидером эпохи была физика, то со временем на авансцену прогресса выдвинулись и другие области знания — биология, экология (особенно социальная), теория систем, синергетика. И конечно же, кибернетика, микроэлектроника, компьютерная техника, информатика. Это последнее обстоятельство имманентно связано с непрерывным возрастанием удельного веса и значения именно информационного

среза как социальных процессов, так и познания природы, ее объектов и свойств, а особенно — человека, его психики, жизни и деятельности. В соответствии с этим в методологической плоскости науки неуклонно повышалась роль *информационного подхода* как особого познавательного инструмента общенаучного масштаба [14–17]. При этом следует подчеркнуть, что эта его функция стала возможной лишь в контексте формирования нетрадиционного типа современных общенаучных средств, важнейшим из которых является системный подход. И именно этот класс новейших теоретико-методологических конструктов в наибольшей степени характеризует научную эвристику наших дней.

Будучи принципиально важным компонентом НТР, компьютерная (информационная) революция вызвала трансформацию типа информатики в обществе — начало перехода от бумажной информатики к безбумажной, электронной [18]. Информатизация общества буквально на глазах нескольких поколений изменила лицо мировой цивилизации: практически иными — информационными (в самом современном смысле слова!) — стали технологии во всех сферах жизни и деятельности человека. Персональный компьютер, микрокалькулятор, цифровой фотоаппарат, мобильный телефон, банкомат, электронный билет на транспорте... — подобных примеров повседневной реальности не перечислить (не говоря уже о сложнейших производствах или научных исследованиях). Понятно, что и производство книги (либо иных полиграфических изделий) не могло остаться в стороне от всеобъемлющего потока электронизации и внедрения микропроцессорной, информационной техники.

ДОСТИЖЕНИЯ МИКРОЭЛЕКТРОНИКИ И КНИГОПЕЧАТАНИЕ

Книга уже давно является важнейшим атрибутом мировой цивилизации: в значительной степени именно в книгах концентрируются для сохранения в социальной памяти достижения человеческого разума в самых различных отраслях жизни общества. Еще в середине XX в. книжная статистика называла впечатляющие, хотя и существенно различные объемы «галактики Гутенберга» — 30, 50 или даже 100 млн *названий* книг, не говоря уже о суммарном тираже [19]. Очень долго именно этот способ хранения и передачи семантической информации в социуме оставался безусловно доминирующим, если не основным. А вместе с тем было ясно, что «роль книги как источника и средства распространения научных сведений ослабляется тем, что для написания книги даже коллективом авторов требуется несколько лет. Издание ее также занимает длительное время. Поэтому многие сведения, сообщаемые в момент опубликования книги, уже имеют давность от трех до пяти лет и часто оказываются устаревшими. В таких быстро растущих отраслях, как физика и химия, книги стареют быстрее; в них используются главным образом книги последних лет издания» [19, с. 85]. Повторим: так было долго, вплоть до свершения информационной революции. Недаром же в свое время появились научные журналы как средство ус-

корения публикации важнейших результатов исследований.

Внедрение новых информационных технологий существенно изменило книгопечатание: резко сократилось время изготовления продукции, повысилось ее качество. Сегодня наиболее ощутимо влияние новейших информационных технологий на допечатную подготовку изданий. Вместе с тем более всего известны распространённые технические средства оперативной полиграфии — ксерокс, принтер и т. п. Отметим, что до 1973 г., когда фирма «Ксерокс» оборудовала компьютер небольшим монитором, электронно-вычислительные устройства практически не затрагивали потребностей печатного дела. Следует также учесть, что для широкого круга потребителей компьютеры были недоступны до 1976 г. Лишь в 1974 г. Стив Джобс и Стефан Возняк начали работу над своим первым компьютером: машину оснастили цветным телевизором, который использовался в качестве дисплея. И только к началу 80-х гг. ЭВМ из грандиозного монстра, некогда занимавшего площадь в 300 м², превратилась в устройство, все еще лишь отдаленно напоминавшее нынешний компьютер.

Появление столь привычной теперь *настольно-издательской системы* связано с именем П. Брайнерда, основателя компании «Алдус Корпорейшн», выдвинувшего идею использования персонального компьютера для создания публикаций в электронной форме, чтобы в дальнейшем воспроизводить их на специальных печатных устройствах. Первая такая система появилась в 1985 г., и с ее рождением традиционные методы набора текстов постепенно отошли в прошлое. До 1984 г. компьютеры в редакциях выполняли функции пишущих машинок, но с созданием лазерного принтера фирмы «Эпл» стало понятно, что распечатывать можно не только текст, но и ретрированные изображения [20].

Работникам издательств постоянно приходится вводить в компьютер большие объемы текстовой информации, опубликованной ранее. Сберегать накопленную годами информацию на бумаге нецелесообразно: на поиск необходимого документа тратится много времени. Подлинным спасением для издательств и редакций периодических изданий стали *системы оптического распознавания текста* — специальные компьютерные программы бесклавиатурного введения информации, способные преобразовывать ее из графического формата в текстовой. Устройством ввода является *сканер*, позволяющий компьютеру «увидеть» текст в графическом формате. Изображение с листа бумаги, помещенного в сканер, превращается в электрические сигналы, в понятную компьютеру последовательность нулей и единиц (т. е. в цифровую форму).

Сегодня структуру традиционного издательства можно разделить на три основных подразделения, у каждого из которых своя роль. Первое из них — редакционное, оно, как и ранее, получает и редактирует рукописи. Отредактированный текст поступает в производственное подразделение, которое разрабатывает оформление будущего издания, верстает его и готовит к передаче в типографию. Маркетинговое

подразделение отвечает за рекламу и продажи подготовленного издания. Внедрение информационных технологий существенно повлияло на разделение труда в издательстве. Ныне издатели корректируют свою редакционную и производственную деятельность так, чтобы конечный продукт легко мог быть приспособлен к потребностям конкретной издательской среды — производства изданий на бумаге, оптических носителях информации или же в Интернете. Это объединяет редакционный и производственный процессы в единый поток [21]. А впереди — новые горизонты.

В течение 1980–1992 гг. японские фирмы-изготовители при поддержке правительства Японии работали над созданием компьютера пятого поколения [22]. Целью проекта было не скачкообразное, как ранее, повышение производительности ЭВМ — основное внимание было сосредоточено на искусственном интеллекте, базой которого выступает логический язык программирования. И именно с этим направлением развития электронной техники связывается новый этап прогресса книгопечатания: внедрение систем искусственного интеллекта открывает перед производством книги невиданные ранее перспективы. Это создание и использование в полиграфии целого комплекса системных решений, в частности, внедрение: систем распознавания компьютером человеческой речи, которое должно повысить производительность труда и снизить зрительное напряжение; систем компьютерного слежения за действиями оператора в реальном времени (например, во время работы на печатной машине), извещения о выявленных ошибках и рекомендациях наборщику, т. е. систем редактирования текста; экспертных систем редактирования, верстки, обработки цветных изображений, цветопроб, поиска повреждений, а также планирования, обучения и т. п., которые позволят операторам с невысокой квалификацией и незначительным опытом выполнять работу высококвалифицированных специалистов; систем защиты электронной цифровой информации от несанкционированного доступа (на основе систем восприятия речи); малогабаритных, удобных и доступных по цене приборов распознавания текста, способных превращать изображение обычных печатных изданий, например, в брайлевский шрифт с тактильным сопровождением либо звуковым воспроизведением, что позволяет незрячим читать или слушать их вместо использования громоздких «брайлевских» изданий [23, с. 206].

В последнее время цифровые средства передачи информации все более активно конкурируют с аналоговыми. Основными преимуществами электронных изданий считаются: мгновенная передача информации потребителю; компактность носителей информации; удобство использования в виде каталогов, энциклопедий, справочников и т. п.; экономия материальных носителей информации — бумаги, картона, пленки и др. Вместе с тем у непрофессионально подготовленных электронных изданий без предварительной редакционной обработки (в основном это неавторитетные источники информации по сравнению с традиционными изданиями) есть и су-

ществленные недостатки: сложность поиска необходимой информации в Интернете из-за ее массовости и неупорядоченности; то, что большая часть информации в интернет-ресурсах не сохраняется навсегда, а может «бесследно» исчезать; сложность соблюдения и сохранения авторских прав; зачастую неудачная подача визуальной информации для чтения (тогда как правила набора и верстки в полиграфии выверены столетиями).

Новейшие достижения в области компьютерной техники и программного обеспечения ощутимо повлияли на восприятие информации человеком и способы ее распространения. Мы являемся свидетелями цифровой революции с присущей ей высокоскоростной передачей информации (гипертекст, графика, анимация, видео- и аудиоинформация). Хотя ныне мы находимся лишь в самом начале цифровой революции: ведь потребность в увеличении информационной емкости и скорости записи средств хранения и передачи информации вновь актуальна.

Автоматизация процесса обработки больших объемов текстовой информации неизмеримо ускорила издательский процесс: авторы представляют материалы в электронной форме и активно пользуются программами проверки правописания (правда, не всегда совершенными), конверторы обеспечивают легкое преобразование одного электронного формата в другой. То, что раньше занимало львиную долю времени у работников издательств — правка и перепечатка текста на бумажных носителях, — сегодня намного быстрее делается на экране монитора компьютера. Верстка текста осуществляется с помощью пакетов специальных программ, а выведение оригинал-макетов, печатание и отделочные работы заменяются генерированием файлов оцифрованной информации на различных носителях.

Важнейшей составляющей издательского маркетинга всегда было размещение рекламы и распространение различных сведений о предлагаемой читателю продукции — книгах, журналах и т. п. Электронное издательство, в частности Интернет, решительно меняет и эту сторону дела. Потребители теперь имеют более удобный и оперативный доступ к информации через системы поиска, виртуальные книжные магазины и форумы, поиск может быть более широким либо узким.

Яркой иллюстрацией этого может служить деятельность британского издательства «Геликон». Основное его достижение — система из целого блока компьютерных баз данных, содержащих около 20 млн слов по различным темам: наука, искусство, история, биографии выдающихся людей, языки, мировая экономика и бизнес, политика, география, спорт и т. п. Поиск осуществляется с базы-каталога, причем имеются гиперсвязи между отдельными элементами баз (текстом, таблицами, иллюстрациями, видео- и аудиоинформацией и т. п.), а их содержание ежедневно обновляется в стандартных языках разметки документов. Информация баз данных используется преимущественно для создания справочно-энциклопедических изданий в виде книг, CD, DVD, а также для размещения в Интернете.

Ныне на рынке электронных изданий наблюдается тенденция возрастания доли онлайн публикаций и сокращения количества изданий на компакт-дисках. В последнее время увеличилось количество названий бумажных изданий — газет, журналов, еженедельников — при одновременном уменьшении тиражей. Тем самым слухи о гибели традиционных печатных изданий и их вытеснении электронными версиями оказались несколько преувеличенными [24; 25, с. 182–225]. Но... прогресс остановить невозможно. И если человечество уверенно шагает к будущему информационному обществу, то традиционные и электронные издания, видимо, еще долго будут сосуществовать в условиях взаимной конкуренции. А объединять их будет вечный поиск все новых и более совершенных форм, в котором реализуется неутолимое человеческое стремление читать, чтобы знать больше, развиваться интеллектуально.

Как видим, уже конец XX в. изменил лицо полиграфии. Что же ожидает книгопечатание в XXI в.? Каковы перспективы его дальнейшего развития?

НОВЕЙШИЕ ПОЛИГРАФИЧЕСКИЕ ТЕХНОЛОГИИ

В области средств передачи информации рубеж тысячелетий, как отмечено выше, был ознаменован прежде всего достижениями цифровой революции (микроэлектронной по своему существу).

Цифровыми, строго говоря, называются способы воспроизведения информации в компьютерных устройствах и системах [18]. Однако цифровую (или, как ее еще называют, дигитальную) печать не следует ставить в один ряд с такими понятиями, как высокий, плоский, глубокий или трафаретный способы печати. Английское *digital* имеет наряду с основным и другое значение — кодированный (преобразованный из аналоговой формы в цифровую), т. е. в данном случае (применительно к печати) имеется в виду способ создания изображения на печатной форме.

Развитие цифровой печати насчитывает около 20 лет, ее пионером был Н. Негропonte, один из основателей медиалаборатории Массачусетского технологического института. Именно он сформулировал принцип получения переменной печатной формы. На рынок цифровой печати первыми со своими предложениями вышли компании, специализирующиеся на научных разработках в области копировальной техники. Многим тогда казалось, что цифровая печатная машина будет увеличенной моделью лазерного принтера. Но уже сейчас на рынке полиграфического оборудования сосуществуют одновременно несколько принципиально различных вариантов цифровых печатных устройств [26, с. 72–76].

При создании таких печатных машин необходимо было решить проблему полного перенесения печатной краски, или тонера, с офсетного цилиндра на бумагу. Первой эту задачу практически решила израильская фирма «Индиго», создавшая в 1993 г. цифровую листовую печатную машину, а в 1995 г. — рулонную. В ксероксах и принтерах роль красителя выполняет сухой тонер, частицы которого достаточно велики, а при попытках уменьшить их размер они

становятся неуправляемыми. Поэтому разработчики цифровых печатных машин создали жидкий тонер — краску, способную удерживать статический заряд, а в основу принципа формирования изображения на запечатываемой поверхности была положена жидкостная электрофотография. Существуют также цифровые печатные машины плоской офсетной печати без увлажнения (сухой офсет).

Что ожидает технологию печати в ближайшем будущем? Производительность оборудования определяется скоростью выполнения самой медленной операции в общем производственном потоке. Когда процессорные устройства станут качественно новыми, проблемы узких мест в системах обработки информации будут менее ощутимы и во всем печатном деле произойдут радикальные перемены. Прежде всего повысится интеллектуальный уровень редакторской, иной издательской и полиграфической управленческой работы на базе компьютерной техники. Компьютеры смогут работать от прикосновения и даже с голоса (звуков речи), комплексно выполняя редакционные, издательские, сугубо печатные, а возможно, и послепечатные операции. Система «человек — компьютер — печатная машина» достигнет более высокого уровня развития, создаст новые возможности для интеллектуального управления технологическими процессами.

В ближайшее время прогнозируется дальнейшее развитие локальных и глобальных компьютерных сетей, создание оперативных систем печати, способных выполнять заказы, переданные через Интернет или спутники связи, усовершенствование допечатных технологий с использованием перенесения информации с компьютера на печатную форму, с компьютера на оттиск и т. п. Расширяются возможности изготовления печатной продукции с использованием гибридных технологий. Наиболее вероятны и дальнейшее повышение производительности печатного оборудования благодаря компьютеризации и автоматизации отдельных операций, увеличение площадей запечатывания с возможностью одновременного печатания в одной машине разными способами печати. Ожидается увеличение красочности, сокращение циклов подготовительных операций, возрастание скорости печати при одновременном уменьшении количества обслуживающего персонала. Продолжается совершенствование отдельных механизмов и деталей машин, распространение обычных, статистических, динамических систем управления производственными потоками, создание новых и улучшение имеющихся материальных носителей информации. Вероятно также появление новых видов печатной продукции, неизвестных сегодня либо находящихся на проектно-изыскательской стадии. Например, известно, что электроника совершила в своем развитии огромный скачок, когда взамен вакуумной трубки был изобретен транзистор, а затем и кремниевый чип, благодаря чему электронные приборы стали более мощными, миниатюрными и более дешевыми.

Популярными ныне становятся нанотехнологии, т. е. производство нанопроductов, к которым относятся любые объекты, имеющие размеры меньше

0,1 мкм (100 нм). Это дает возможность значительно экономить материалы и энергию. Для компьютерного оборудования, используемого теперь в полиграфии, это означает увеличение скорости процессов и повышение плотности записи информации.

Следует впрочем заметить, что научно-технический прогресс не всегда идет строго по восходящей — от низшего к высшему, от простого к более сложному и совершенному. Иногда значительное распространение получают технические средства, о существовании которых через какое-то время человек даже не может вспомнить. Так произошло, например, с микрофильмами и микрофишами. Время доказало — они не оправдали возлагавшихся на них надежд из-за сложности в использовании и очень высоких затрат на их внедрение. Когда-то считалось, что традиционную печать смогут заменить именно микроиздания — микрофильмы и микрофиши. Микрофильм тогда снимался на обычную киноленту: на одном кадре размещали разворот книги (две страницы), используя при этом и негативную, и позитивную пленки. В 80-е гг. прогнозировали, что в конце XX в. 95% всех документов будут храниться в виде микрокопий, поскольку уже к тому времени вес технической документации на какое-то изделие приближался к весу самой продукции (например, самолет «Боинг-707» весил 26 т, а чертежи и инструкции к нему — 23 т) [27, с. 131; 28]. Однако это предвидение не осуществилось, хотя в свое время благодаря микроизданиям и удалось сэкономить свыше 90% площади книгохранилищ. Скажем, используя микрофильмы, в одном шкафу можно было разместить и книгохранилище, и каталог к нему. Некоторые библиотеки уже тогда отказались от хранения газет и других периодических изданий в обычном виде, переведя их на микрофильмы. В то же время микрофильмирование имело и существенный недостаток: при чтении неудобно было протягивать длинную пленку через специальный проекционный аппарат, особенно когда нужная информация содержалась где-то в середине, допустим, двухсотстраничной книги (длина микрофильма с книги такого объема составляла около 4 м).

Микрофиши (называемые еще микрокартами) в сравнении с микрофильмами были более совершенными. Размер самых распространенных микрокарт составлял 10,5×14,5 см, толщина — несколько десятых миллиметра. Книга объемом до трехсот страниц могла быть заменена 4–10 микрокартами толщиной не более 1 мм. Тридцать томов третьего издания Большой Советской Энциклопедии (1969–1978), весившие около 100 кг, в микрофишах имели вес до 1 кг. Микрофильмирование давало возможность каждому при необходимости заказать новую книгу и получить ее из книгохранилища в почтовом пакете. Впрочем, следует отметить, что даже в идеальных условиях микроиздания можно было хранить лишь 25–30 лет, а не столетия, как летописи или печатные издания.

Подобных примеров, касающихся других научных и конструкторских разработок в области печати и хранения информации, можно привести немало. Так было, в частности, с механическим и полуавто-

матическим фотонабором — они были востребованными до тех пор, пока не появилась электроника. Это же касается и наборных машин — линотипов и монотипов. Нечто подобное произошло и с печатанием газет на расстоянии в специально оборудованных автоматических киосках, размещенных на далеком от редакций расстоянии. Появление компьютера и Интернета в корне изменило принципы печатания и передачи единичных экземпляров газет на расстояние.

Как видим, научно-технический прогресс наряду с широкими основными магистралями неизбежно имеет, образно говоря, и узкие переулки, и тупиковые ветви. В частности, в области книгопечатания ныне это достаточно очевидно. Будущее, несомненно, может принести новые сюрпризы и зигзаги в целом непростого маршрута. Вместе с тем несомненным в эпоху информационной революции является то, что магистральный путь дальнейшего совершенствования полиграфических технологий имманентно связан с развитием микроэлектроники [18]. Цифровая революция углубляется и одновременно расширяет свое русло в социуме.

ЗНАЧЕНИЕ ПОЛИГРАФИИ В СОВРЕМЕННОЙ КУЛЬТУРЕ

Полстолетия уже пролетело с тех пор, как канадский философ и социолог М. Маклюэн выступил с широковещательным прогнозом о конце книги. «Галактике Гутенберга» он прямо противопоставил «электронную галактику» [29]. Тогда, в 1962 г., эпоха компьютерной техники лишь зарождалась (изобретение микропроцессора было еще впереди!), и потому большинству читателей брошенный вызов казался особенно дерзким, совершенно неоправданным. (Ведь не случайно и сам этот вызов тогда мог быть оформлен лишь в виде *книги!*) Сегодня же, через пять десятилетий, об этой проблеме можно говорить значительно спокойнее, более взвешенно и обоснованно.

Жизнь доказала, что конец «галактики Гутенберга» не настал — и сейчас мы говорим о печатной индустрии, о полиграфии как о живом, развивающемся организме. Другое дело, что перспективы книги и всего печатного дела (что гораздо шире) органично связаны с прогрессом электроники.

Печатное слово доказало не только свою жизнеспособность в течение многих столетий, но и особое значение для судеб цивилизации. Ведь его основная функция — надежная передача информации в пространстве и времени, а это уже давно важнейший для социума процесс. В XX в. наука вплотную подошла к раскрытию тайны того, что уже давно интуитивно ощущалось в обществе: наряду с веществом и энергией информация необходима не только для жизни и деятельности человека, но и для всего живого в природе [6, 30]. Как отмечалось ранее, именно это положило начало формированию информационного подхода как особого компонента современной общенаучной методологии. Что же касается истории общества, то переход к бумажной информатике (вначале к письменности, а затем и к книгопечатанию) означал заметный скачок в развитии мировой цивилизации.

С ним имманентно связана вся дальнейшая культура человечества — как духовная, так и материальная.

Оглядываясь в прошлое и перебрасывая мостик из наших дней в будущее, следует в каждой из этих эпох особо отметить исключительно большое значение книги для таких отраслей общественной жизни, как образование и воспитание новых поколений, наука, техника (и сфера технологий), медицина и здравоохранение, строительство, транспорт, связь и коммуникации. Впрочем, к ним следует добавить и другие, не менее важные стороны развития человечества — религию, мораль, философию, искусство, право. И все равно этот перечень не будет полным. Как видим, книга имеет прямое и непосредственное отношение ко всем без исключения сторонам жизни человека и общества. По существу это вся многоцветная палитра исторического развития социума.

В течение многих десятилетий и даже столетий уровень общей культуры человека далеко не в последнюю очередь определялся кругом его чтения (т. е. в конечном счете — воздействием книги на его мировосприятие и формирование мировоззрения). Сегодня эта своего рода монополия книги отошла в прошлое, уступив место гораздо более мультиплицированному социокультурному влиянию новых видов источников информации.

Ныне социологи, книговеды и культурологи отмечают как всеобщую тенденцию ощутимое снижение интереса широких масс к чтению (что связывается с такими культурными факторами, как прогресс радио и телевидения, развитие видеотехники и компьютерных сетей, прежде всего Интернета). К сожалению, резко сузился круг чтения молодежи, в том числе учащихся средней школы и студентов. Но и в этом плане не все так однозначно, как может показаться на первый взгляд: чтение книг остается серьезной проблемой использования свободного времени членами общества, и все еще имеется немало людей, для которых оно является стратегией жизни [31, 32].

Безусловным стал тот факт, что человечество неуклонно приближается к реалиям информационного общества с его новейшими электронными технологиями, и уже это обстоятельство не может не потеснить позиции книги в бюджете свободного времени нашего современника. Но вместе с тем в ходе углубления этой проблемы возникли такие «гибридные», переходные инструменты культуры, как электронная книга, электронный учебник, электронная библиотека. Их появление свидетельствует о том, что общество постоянно ищет новые формы, способные совместить привычные достоинства книги (которые мировая цивилизация оценила уже давно) с сильными сторонами новейшей информационной техники.

Вся история человечества доказывает, что книга — отнюдь не рядовая форма бумажного представления информации, не просто один из многих ее источников в обществе. Это громадное культурное обретение социума, связанное не только с материальным воплощением социальной памяти многих поколений, но и с удовлетворением разнообразных духовных потребностей человека — познавательных, коммуникационных, эстетических, этических. Разными своими

гранями книга оказывается сопричастной к выявлению деятельностной сущности человека, к удовлетворению его внутренней потребности в самовыражении и творчестве. Она — важнейший способ приобщения личности к духовной культуре человечества.

Любить книгу в ее традиционной форме и привычный процесс чтения призывают человека многие факторы психологического характера. Когда-то академик И. Т. Фролов, один из исследователей информационной революции [33], говорил, что не может представить себе даже далекого потомка, которому не хотелось бы сесть вечером у любимой зеленой лампы, взять с полки томик «Анны Карениной» и забыть обо всем, читая (и возможно, не в первый раз) строки Л. Н. Толстого... В самом деле, чтение произведений художественной литературы — уникальное, ничем не заменимое удовольствие, дающее вместе с тем возможность глубже познавать мир «в образах и картинах». Но нет необходимости доказывать, что это лишь одна из множества граней социальной роли книги. С нею, например, органично связаны все основные религии мира (как известно, название «Библия» происходит от греческого *biblion* — книга): именно священные книги издревле определяют общую семантику и фундаментальные основы любого религиозного мировоззрения. Закономерны также огромные тиражи всевозможных учебников, пособий, справочников, словарей и различных учебно-методических изданий: они предназначены для широчайшей аудитории тех, кто учится и, с другой стороны, преподает, учит других. При этом отметим, что во многих странах мира образовательная идеология все больше руководствуется принципом «обучение в течение всей жизни», и это имеет особое значение для современности [34].

В качестве примеров мы назвали лишь несколько областей развития общества, где роль книги всегда была ощутимо велика и со временем отнюдь не снижается. Но этот перечень, конечно, можно продолжать достаточно долго. Особое место в нем, естественно, следует отвести прогрессу науки и техники, значению книги в развивающейся системе научно-технической информации, но совершенно очевидно, что именно эта проблематика читателям сборника «НТИ» знакома глубже всего. И все же к этому важнейшему вопросу мы еще вернемся в последнем разделе статьи (здесь ограничимся лишь упоминанием).

Учитывая массовый характер использования книги в обществе, обратим внимание на психологическую окрашенность и тональность этого процесса. Принципиально важной особенностью книги уже давно стала ее психологическая совместимость с человеком: для подавляющего большинства членов общества чтение книги (даже работа с ней, когда речь идет о специальных изданиях) является процессом довольно комфортным, не требующим высокого нервного напряжения, нередко даже приятным, во всяком случае привычным. Важно уже то, что книгу в любой момент можно отложить, а затем вернуться к ней, когда будет удобно. Приметная особенность и в том, что текст книги можно перечитывать неодно-

кратно, причем разные места — в любой последовательности. Всегда есть твердая уверенность, что с оставленным, отложенным текстом ничего не случится, он не исчезнет (как бывает в компьютере) — в этом отношении человек совершенно спокоен. Такие психологические характеристики уже в значительной степени способны объяснить спокойный взгляд большинства книговедов и культурологов на перспективы книжного дела, по крайней мере, на период жизни нынешних поколений.

Вместе с тем в современной истории общества есть и другие причины, позволяющие смотреть на будущее книги с оптимизмом. К ним, например, относится «революция досуга», наблюдавшаяся в Европе в последние десятилетия XX в. [32, с. 25–32]. Наглядной иллюстрацией ее может служить ситуация в ФРГ, по крайней мере на ее «старых землях». В 50-е гг. там была шестидневная 48-часовая рабочая неделя, календарный год насчитывал 279 рабочих и 86 нерабочих дней. Изменения в распределении годового бюджета времени наблюдались уже в 70-е гг.: рабочая неделя стала пятидневной и 42-часовой, количество рабочих дней уменьшилось до 238, а выходных, праздничных и отпускных — возросло до 127. В начале 90-х рабочая неделя, оставшись пятидневной, сократилась до 38 часов. Количество рабочих дней не изменилось, а суммарное свободное время заметно возросло. Подобные процессы наблюдались и в других странах, что и дало основание говорить о «революции досуга». Она, несомненно, связана и с тем, что больше времени может уделяться чтению (среди других способов использования досуга). Такой была тенденция в конце XX и начале нынешнего века.

Конечно, в наши дни реалии мирового финансово-экономического кризиса заставляют вносить существенные коррективы в оценку положения на рынке труда, иначе видеть социокультурные проблемы и перспективы, строить иные прогнозы на будущее. Но в то же время в масштабах человеческой истории общая мировая тенденция постепенного снижения количества общественно-необходимого рабочего времени и повышения доли свободного времени в жизни людей в целом представляется все-таки правомерной. И это, думается, один из весомых позитивных факторов и социального развития вообще, и понимания роли книги в будущем человечества. Во всяком случае, это явно аргумент в пользу оптимистической оценки дальнейшей судьбы книги.

В заключение данного раздела подчеркнем еще один принципиально важный момент: значение полиграфии в мировой культуре не ограничивается лишь печатанием книг (а также журналов и газет). Полиграфическая продукция давно уже стала многообразной — плакаты и календари, открытки и конверты, инструкции к различным товарам и их упаковка, этикетки, ценные бумаги, деньги и знаки почтовой оплаты, картографическая и акцидентная продукция и т. п. В эпоху НТР, когда и общему качеству, и эстетическому оформлению любой продукции уделяется немало внимания, эта сторона проблемы также постоянно должна быть в поле зрения исследователей полиграфии как многогранного фе-

номена. Ведь это понятие уже давно обозначает не только совокупность особого рода технических устройств и технологий, но и отрасль экономики, промышленности, охватывающую производство всех видов печатной продукции.

Исторически книгопечатание было первой функцией полиграфии. И сегодня оно остаётся важнейшей, но уже далеко не единственной задачей этого сложного, комплексного социального организма, весьма заметного компонента современной культуры человечества. Полиграфия — во всем объеме данного понятия — постоянно и многосторонне ощущает на себе воздействие импульсов информационной революции.

ИНФОРМАТИЗАЦИЯ ОБЩЕСТВА И БУДУЩЕЕ КНИГИ

Уже в течение ряда десятилетий информационная плоскость НТР неизменно и непрерывно привлекает к себе внимание исследователей. Многие авторы давно поспешили декларировать переход к информационному обществу, т. е. полное торжество его принципов. На наш взгляд, это не совсем корректная позиция. Точнее, видимо, говорить о тенденции и начавшемся переходе к этому новому этапу в развитии общества, качественно иному, нежели ранее, состоянию социума.

Для того чтобы разобраться в существе вопроса, очевидно, необходимы более или менее определенные критерии. Один из них (имеющий сугубо качественную природу, без малейших количественных признаков) — ориентация общества на принципы информационного подхода, вовлечение понятия информации в ядро мировоззренческих категорий. Другой (также качественного характера) — осознание информации как важнейшего социокультурного фактора эпохи, многогранно детерминирующего жизнь отдельного человека и социума в целом. Философская и научно-техническая литература второй половины прошлого века не оставляет сомнений в том, что характеристики такого рода давно уже достигнуты развитием общества. Гораздо сложнее нащупать более или менее четкие критерии количественной природы. А. Д. Урсулу удалось выйти, по крайней мере, на один из них. Опираясь на работы видных экономистов и социологов, он сформулировал положение о том, что общество может считаться информационным лишь при условии, что почти 90% его населения тем или иным образом занято в сфере производства информационных продуктов, а не в промышленности или сельском хозяйстве [35, с. 354]. Сегодня, конечно, до этого далеко даже наиболее развитым странам, пионерам и лидерам информатики. И это означает, что достижение рубежей информационного общества — пока еще дело будущего. Но движение в этом направлении начато и, безусловно, усиливается со временем.

Вот почему правильнее считать, что современное состояние общества — переходное, это постепенный переход от индустриальной цивилизации к информационной. И необходимым условием этого перехода является процесс *информатизации общества*. В от-

личие от компьютеризации или электронизации (явлений сугубо технических либо, самое большее, технико-экономических) информатизация общества представляет собой сложный феномен общесоциальной природы, «процесс, в котором социальные, технологические, экономические, политические и культурные механизмы не просто связаны, а буквально сплавлены, слиты воедино» [8, с. 34]. Сущность информатизации — качественное преобразование (на базе новейших научно-технических, технологических, организационных и иных достижений человечества) всей информационной инфраструктуры жизни общества с целью оптимизации результатов социально значимой деятельности любого рода.

Совершенно очевидно, что информатизация общества органично и вместе с тем многообразно связана с развитием полиграфии и средств массовой информации и коммуникации. В этой отрасли, кстати, очень отчетливо видно, как ускоряется со временем научно-технический и вообще социокультурный прогресс. Тысячелетия отделяют рисунки первобытных людей на стенах пещер и первые оттиски на глиняных табличках от печатания на станке с металлических печатных форм. Лишь сотни лет лежат между традиционным книгопечатанием и возможностью вывода печатного текста из памяти компьютера с помощью лазера. И только десятилетия отделяют беспроводное радио и телефон от появления Интернета и сотовой связи. Ручной набор господствовал в печатном деле почти 500 лет, строкоотливные и буквоотливные наборные машины обеспечивали процессы набора и верстки в течение 100 лет, фотонабор использовался примерно 25 лет. Новые технологии всегда вытеснялись новейшими, более прогрессивными.

Еще в 1980 г. известный специалист издательского дела Ф. Романо высказал интересную мысль о необходимости соединить набор с печатанием [20, с. 16–19]. Уже в конце XX в. фирма «Ксерокс» начала выпуск устройств, соединивших компьютерный набор и печать.

С помощью компьютера авторы получили возможность непосредственно вводить свои произведения в его память. Ныне компьютеры стоят на рабочих местах сотрудников издательств, помогая не только автору, редактору, корректору, но и иллюстратору книги — художнику. Особо большую роль электроника играет в процессах цветной репродукции, где раньше использовали ручной труд, а затем — фототехнику.

Несмотря на то, что сейчас немало говорят и пишут о новой физической форме книги, в ближайшем будущем традиционной ее форме вряд ли что-либо угрожает: ведь новые формы и традиционная книга — не конкуренты-антагонисты, а дополняющие друг друга средства распространения информации. Сегодня с печатными изданиями конкурируют электронные. Их развитие осуществляется в двух направлениях — техническом (обеспечивающем совершенствование средств изготовления и распространения) и программном, задачами которого являются автоматизация процессов разработки, со-

вершенствование практического использования и дизайна электронных изданий.

Известно, что появление нового средства коммуникации не приводит к вытеснению медийных инструментов, использовавшихся человеком ранее. С появлением книгопечатания люди не перестали писать от руки; когда стала доступной радиосвязь, газетное производство не исчезло; появление телевидения не привело к гибели радиовещания. Старые средства передачи информации теряют былое значение, но сохраняются и совершенствуются. То же самое происходит с книгой.

Естественно, что сегодня в книжном деле можно наблюдать новаторские, оригинальные проекты. Так, например, британское издательство «Харпер Коллинз» совместно с компанией «Пенвизард» и печатным центром «ХП Индиго» решили проблему издания персональных книг на заказ. Разработано необходимое программное обеспечение, позволяющее любому пользователю Интернет-сети в интерактивном режиме придавать нарисованному на компьютере в издательстве персонажу черты заказчика издания либо его друзей, сотрудников, членов семьи, а в текст вставлять их имена. Таким образом получается издание, героями которого становятся персонажи заказчика.

В процессе практической реализации такого издания заказчик со своего компьютера обращается на соответствующий сайт, заполняет там анкету-шаблон издательского центра «ХП Индиго», где размещается электронная версия подготовленного издательством оригинального произведения. Центр осуществляет перевоплощение оригинальных издательских персонажей в образы с заказанными чертами, через Интернет согласует версию с заказчиком и после исправлений (при необходимости) почтой высылает полноцветное персональное издание на бумаге либо в электронной форме. Такие издания особо популярны среди детей и их родителей, но могут также быть прекрасной рекламой и фирм, и новых технологий, и аппаратно-программных комплексов, и новой техники. Это уже новая издательско-полиграфическая технология по принципу «с монитора в печать» [36].

В Российской Федерации аналогичный проект реализует издательство «Вебов и Книгин»: заказчику предлагается издать многокрасочную книгу на бумаге или электронную (на компакт-диске), в которой главным героем и по тексту, и в иллюстрациях будет его ребенок. Издательство подготовило несколько сказок, написанных специально для этого проекта известными писателями и оформленных лучшими художниками [37].

В последнее время все шире применяют технологии виртуальной реальности, основанные на компьютерной трехмерной графике и создании иллюзии реальности. Этот способ распространения информации особенно продуктивен для представления новых образцов техники, рекламы музеев и выставок, презентации архивов и т. п.

В современных условиях по-прежнему актуальна разработка систем совершенного машинного перевода с одного языка на другой. В этой области существ-

уют два основных направления. Первое: на новой технологической основе развивается традиционная словарная база, документируются языковые структуры и результаты воплощаются в новые электронные словари, причем издание классических бумажных словарей значительно облегчается и совершенствуется благодаря возможности готовить их с помощью компьютерной техники. Второе: использование методики Ж. Ф. Шампольона, который расшифровывал тексты иероглифической древнеегипетской письменности путём сравнительного анализа (подобного тому, как составляются практические двуязычные разговорники с готовыми фразами на все случаи жизни). Понятно, что оба эти направления многогранно используют результаты информатизации общества.

Особо отметим специфику научно-технической книги и ее перспективы на будущее. Как неоднократно подчеркивалось ранее, распространение научно-технической информации с помощью традиционных бумажных средств – книг и журналов – имеет существенные недостатки [19, 38, 39]. Во-первых, имеется большой разрыв во времени между окончанием исследования и публикацией его результатов. Во-вторых, из года в год становится все труднее свободно ориентироваться в свежей книжной и журнальной продукции. В 80-е гг. XX в. было подсчитано, что после завершения новой научной разработки информация о ней в журнале появлялась через один – два года, в монографии – через три. Сегодня эти сроки сократились в 2–3 раза, но сама проблема оперативного ознакомления общества с полученными результатами научно-исследовательской и инженерно-конструкторской работы не решена до сей поры. Конечно, основным глобальным источником информации стал Интернет, содержащий уже миллионы веб-страниц. Но Интернет имеет свои проблемы: прежде всего это комплекс вопросов, связанных с авторским правом и безопасностью, вынуждающих, в частности, прибегать к системе закрытых сайтов, что само по себе означает заметный недостаток глобальной информационной сети. Очевидно, на ближайшее будущее принципиальное решение проблемы лежит в плоскости адекватного сочетания традиционных (бумажных) и электронных средств, оптимизации их взаимодействия.

В этом плане накоплен уже определенный опыт. Упомянутое ранее издательство «Геликон», например, выпустило 600-страничный полноформатный печатный указатель веб-сайтов, содержащий резюме более 5000 тщательно отобранных электронных хранилищ информации по 14 тематическим разделам знания (природа, искусство, история, литература, медицина, музыка, география, политика, наука, общество, спорт, техника, философия, экология). К книге бесплатно прилагается её электронная версия.

Видимо, именно этот путь — сочетание достоинств и преимуществ традиционных (бумажных) и электронных носителей информации — и следует признать стратегическим курсом будущего развития книги вообще, любого вида и жанра (а не только научно-технической).

Современная конструкция книги, зародившись в Древнем Риме, существует уже более полутора тысячелетий. Неудивительно, что сегодня она претерпевает определенные изменения. Еще большие трансформации, несомненно, предстоят в будущем: впереди долгая эпоха сосуществования, активного взаимодействия и вместе с тем соревнования традиционной книги с электронными средствами информатики. «Все течет, все изменяется» – эти слова древнегреческого мыслителя Гераклита, безусловно, относятся и к книге. Вот лишь конца «галактики Гутенберга» ожидать не следует. Книга – слишком важное достижение человечества на пути совершенствования механизмов социальной памяти и духовной жизни общества, чтобы с нею можно было так легко и безнаказанно распрощаться, попросту сдав ее в архив истории.

Информационная революция совершила подлинный переворот в производстве и использовании книги, в полиграфии вообще как особой и весьма важной области развития мировой культуры. Трансформация функций и возможностей книги, несомненно, меняет и лицо современной библиотеки – в ней тоже появилась микроэлектронная техника. Этот комплекс вопросов принадлежит к научно-методологическим проблемам не только ряда социально-гуманитарных отраслей знания (книговедения, культурологии, библиотекведения и др.), но и, конечно, полиграфии и информатики.

СПИСОК ЛИТЕРАТУРЫ

1. Современная научно-техническая революция. Историческое исследование. – М.: Наука, 1970. – 256 с.
2. Марков Н. В. Научно-техническая революция : анализ, перспективы, последствия. – М.: Политиздат, 1973. – 240 с.
3. Научно-техническая революция. Общетеоретические проблемы. – М.: Наука, 1976. – 206 с.
4. Готт В. С., Семенюк Э. П., Урсул А. Д. Научно-техническая революция и современная наука // Философские науки. – 1981. – № 1. – С. 3–12.
5. Научно-техническая революция и развитие научного познания. – Вып. 1. – Баку: Элм, 1985. – 212 с.
6. Винер Н. Кибернетика, или управление и связь в животном и машине : пер. с англ. – 2-е изд. – М.: Сов. радио, 1968. – 326 с.
7. Соболев С. Л., Китов А. И., Ляпунов А. А. Основные черты кибернетики // Вопросы философии. – 1955. – № 4. – С. 136–148.
8. Ракитов А. И. Философия компьютерной революции. – М.: Политиздат, 1991. – 288 с.
9. Семенюк Э. П. Технологический этап научно-технической революции и информатика // Научно-техническая информация. Сер. 1. – 1995. – № 1. – С. 1–9.
10. Семенюк Э. П. Информатика и современные проблемы философии науки и техники // Научно-

- техническая информация. Сер. 1. – 2005. – № 1. – С. 1–10.
11. Лю Ган. Философия информации и основы будущей китайской философии науки и техники // Вопросы философии. – 2007. – № 5. – С. 45–57.
 12. Урсул А. Д. Информатизация общества. Введение в социальную информатику. – М.: АОН, 1990. – 192 с.
 13. Семенюк Э. П. Информатика и современный мир. Философские аспекты. – Львов: Укр. акад. печати, 2009. – 283 с.
 14. Семенюк Э. П. Информационный подход к познанию действительности. – Киев: Наук. думка, 1988. – 240 с.
 15. Информационный подход в междисциплинарной перспективе (материалы «круглого стола») // Вопросы философии. – 2010. – № 2. – С. 84–112.
 16. Урсул А. Д. Информация и информационный подход: от информатики к глобалистике // Научно-техническая информация. Сер. 1. – 2012. – № 2. – С. 1–11.
 17. Семенюк Э. П. Информационный подход и реалии современности // Научно-техническая информация. Сер. 1. – 2013. – № 1. – С. 1–11.
 18. Глушков В. М. Основы безбумажной информатики. – М.: Наука, 1982. – 552 с.
 19. Михайлов А. И., Черный А. И., Гиляревский Р. С. Основы информатики. – М.: Наука, 1968. – 756 с.
 20. Романо Ф. Принт-медиа бизнес : пер. с англ. – М.: Принт-Медиа центр, 2006. – 456 с.
 21. Гехман Ч. Рабочий поток : пер. с англ. – М.: Москов. гос. ун-т печати, 2004. – 252 с.
 22. ЭВМ пятого поколения. Концепции, проблемы, перспективы : пер. с англ. – М.: Финансы и статистика, 1984. – 110 с.
 23. Розум О.Ф., Величко О.М., Мельников О.В. Таємниці друкарства : Минуле. Сучасне. Майбутнє. – 2-е вид., перероб. і доп. – Львів: Укр. акад. друкарства, 2012. – 280 с.
 24. Карьер Ж.-К., Эко У. Не надейтесь избавиться от книги! Интервью Ж.-Ф. де Тоннака : пер. с фр. – СПб.: Симпозиум, 2010. – 336 с.
 25. Дурняк БВ, Штангрет АМ, Мельников О.В. Видавничо-поліграфічна галузь України: Стан, проблеми, тенденції. Статистично-графічний огляд. – Львів: Укр. акад. друкарства, 2006. – 274 с.
 26. Ситников В. П. Издательское дело : Основы. История. Взаимосвязь техники и технологии. – М.: Слово ; АСТ, 2002. – 223 с.
 27. Розум О. Ф. Таємниці друкарства. — Київ: Техніка, 1980. — 143 с.
 28. Иванов Р. Н. Репрография. – 2-е изд., перераб. и доп. – М.: Экономика, 1986. – 336 с.
 29. Маклюэн М. Галактика Гутенберга. Становление человека печатающего : пер с англ. – М.: Академ. проект, 2005. – 496 с.

30. Винер Н. Кибернетика и общество : пер. с англ. – М.: Иностр. лит-ра, 1958. – 200 с.
31. Чтение как стратегия жизни: мат. Междунар. науч.-практ. конф., Москва, 14 дек. 2006 г. / сост.: М. А. Ермолаева, Ю. П. Мелентьева, А. Ю. Самарин. – М.: Наука, 2006. – 220 с.
32. Сеньківський В. М., Олянишен Т. В., Штангрет А. М., Мельников О. В. Читання як проблема вільного часу : За результатами соціолого-статистичних досліджень. – Львів : Укр. акад. друкарства, 2009. – 160 с.
33. Моисеев Н. Н., Фролов И. Т. Высокое соприкосновение. Общество, человек и природа в век микроэлектроники, информатики и биотехнологии // Вопросы философии. – 1984. – № 9. – С. 24–41.
34. Энгельштад К. Интеллектуальное общество, информационная и коммуникационная технология и обучение в течение жизни // Междунар. форум по информ. – 2001. – Т. 26. – № 1. – С. 11–14.
35. Урсул А. Д. Переход России к устойчивому развитию. Ноосферная стратегия. – М.: Издат. дом «Ноосфера», 1998. – 500 с.
36. HarperCollins Publishers. – URL: <http://www.harpercollins.com/>
37. Интернет-издательство «Вебов и Книгин». — URL: http://www.wandk.ru/?page_id=60
38. Арский Ю.М., Гиляревский Р.С., Туров И.С., Черный А. И. Инфосфера: информационные структуры, системы и процессы в науке и обществе. – М.: ВИНТИ, 1996. — 489 с.
39. Иванов Д. В. Виртуализация общества : Версия 2.0. – СПб: Петербургское востоковедение, 2002. – 224 с.

Материал поступил в редакцию 17.09.13.

Сведения об авторах

МЕЛЬНИКОВ Александр Валериевич – кандидат технических наук, директор издательства Украинской академии печати, г. Львов
E-mail: melnikov_alex@mail.ru

СЕМЕНЮК Эдуард Павлович – доктор философских наук, профессор, академик Украинской академии информатики и Международной академии ноосферы (устойчивого развития), член Нью-Йоркской академии наук, заведующий кафедрой философии Национального лесотехнического университета Украины, г. Львов
E-mail: petro_mak@ukr.net

ОРГАНИЗАЦИЯ ИНФОРМАЦИОННОЙ РАБОТЫ

УДК 004.7.03:(048)

В. М. Голубев, Е. Б. Дудин, В. Н. Ушаков

Систематизация, особенности развития и приложение технологии Грид-систем (Обзор)

Предлагается систематизация современных Грид-систем различных типов по основным категориям и признакам, определяющим их технические и функциональные возможности. Рассматриваются особенности интеллектуальных Грид-систем. Перечислены направления, по которым развиваются современные модели, алгоритмы и инструментальные средства для разработки программного обеспечения Грид-систем, в том числе, с элементами искусственного интеллекта. Приводятся сведения о реализованных интеллектуальных программно-аппаратных системах мониторинга и оценки технического состояния объектов в национальных и региональных сетях магистрального электроснабжения

Ключевые слова: Грид-системы, распределенные системы, искусственный интеллект, ПО, алгоритмы, облачные вычисления, интеллектуальные Грид-системы, архитектуры, сервисно-ориентированный подход, энергетические системы

ВВЕДЕНИЕ

Технология Грид является способом организации сетевых вычислений, ориентированным на виртуализацию, когда части задачи распределяются по всем свободным в данный момент, часто географически разнесённым, ресурсам сети [1-6]. Это одно из перспективных и успешно развивающихся направлений современных технологий разработки и использования вычислительных сетей и программного обеспечения. Технология Грид – это распределенная программно-аппаратная компьютерная среда, с принципиально новой организацией вычислений и управления потоками знаний и данных. Инфраструктура Грид позволяет интегрировать большой объем географически удаленных компьютерных ресурсов. Распределённые Грид-вычисления с использованием «решёток» вычислительных ресурсов реализуются преимущественно в архитектурах глобальных, региональных и учрежденческих компьютерных сетей.

Технология Грид впервые была разработана и применена в европейском ядерном центре ЦЕРН (CERN) (Conseil Europeen pour la Recherche Nucleaire – Европейский совет по ядерным исследованиям), а позднее широко распространилась по всему миру [2]. Примерами крупных проектов по созданию Грид-систем являются TeraGrid (США) и DataGrid (ЕС) [7]. С внедрением современных «облачных» технологий вычислений и программирования [1, с.1074-1082; 8-18], а также распределенного хранения и обработки

данных глобальная сеть Интернет становится основной программно-аппаратной платформой для массового использования распределенных вычислительных ресурсов Грид-систем различных типов. Кроме того, при коммуникации электронных устройств с поддержкой функций самонастройки и самообучения, т. е. базовых функций искусственного интеллекта, образуются Грид-системы нового вида – интеллектуальные сети (Smart Grid) [1, с. 928-1045, 1098-1115; 3; 18-21], главная особенность которых заключается в поддержке безотказной коммуникации и более эффективного распределения вычислительных ресурсов.

СИСТЕМАТИЗАЦИЯ ГРИД-СИСТЕМ

С целью более чёткого сопоставления характеристик распределенных Грид-систем различных типов их можно систематизировать по следующим основным категориям [1, с. 928-1115; 2; 22-28]:

- масштаб географического охвата (космические [29, 30], глобальные [7, 29], национальные [1, с. 928-976, 998-1027; 6], региональные [1, с. 928-976, 998-1027; 4-5], домашние [20], учрежденческие [5] системы);
- назначение (промышленные [1, с. 952-956, 977-997, 998-1027, 1074-1082], военные [31-33], коммерческие [34-38], учебные [39, 40]);
- тип линий связи (проводные с выделенными слаботочными линиями [1, с. 928-951, 952-956, 977-997, 998-1027], проводные с передачей по силовым кабе-

лям электроснабжения [1, с. 1074-1082], беспроводные [1, с. 928-951, 977-997; 3; 39-47]);

- способ управления (с централизованным управлением; одноранговые, децентрализованные или пиринговые сети (от англ. peer-to-peer, P2P)) [26];
- уровень программного обеспечения (сетевые сервисы и «облачные вычисления» [34-36; 38], программы промежуточного уровня [22,23], системы виртуальных серверов [37]);
- вид передачи и обработки данных (с асинхронной обработкой или с синхронной в режиме реального времени [1, с. 952-956, 998-1027]);
- количество одновременно подключенных пользователей в сети (от нескольких узлов до многих миллионов пользователей [25]);
- топологическая сложность (одно- или много-ранговые «решётки» [33], интеллектуальные «микрорешётки» ресурсов [1, с. 1074-1082]);
- показатели надёжности и защищённости от программно-аппаратных отказов [1, с. 998-1027; 19; 21];
- энергетическая эффективность [1, с. 1074-1082];
- степень реализации функций искусственного интеллекта (простые, самонастраиваемые, самообучаемые) [1, с. 928-1045, 1098-1115; 18].

В принципе, представленный краткий перечень категорий позволяет определить общие технические и функциональные характеристики практически для любых Грид-систем общего и специального применения.

Независимо от конкретных технических характеристик, все системы различных типов, характеризующиеся перечисленными категориями и показателями свойств, объединяются общим признаком. Все они создаются с использованием Грид-технологии, т.е. на основе способа организации сетевых вычислений, ориентированного на виртуализацию, когда части решаемой вычислительной или коммуникационной задачи распределяются по всем свободным в данный момент (возможно, географически разнесённым) ресурсам сети.

БАЗОВЫЕ КОНЦЕПЦИИ И АРХИТЕКТУРНЫЕ ОСОБЕННОСТИ ИНТЕЛЛЕКТУАЛЬНЫХ ГРИД-СИСТЕМ

Современное общество базируется на распределительных сетях электроэнергии, газа, нефти, связи. Более эффективное, надежное управление этими сетями - важная задача, на решение которой направлено большое количество ресурсов. Предполагается, что в США к 2030 году на эту задачу будет выделено более 560 млрд долл. Важная цель интеллектуальных систем Грид – обеспечить обратную связь в контуре управления на основе динамической информации об объемах потребления или параметрах сети для оптимизации потребления ресурсов во время пиковых нагрузок. Эта концепция сходна с подходом, используемым телекоммуникационными компаниями для сглаживания спроса на услуги во времени. В США переход к интеллектуальным Грид-системам

активно развивается на базе государственных и частных инвестиций. В соответствии с принятым в 2009 году Конгрессом США законом ARRA (The American Recovery and Reinvestment Act) выделено 4,5 млрд долл. на гранты для поддержки производства, закупки и реализации интеллектуальных технологий. Значительные усилия в данном направлении предпринимают Австралия, Европейский Союз, Южная Корея, Япония и многие другие страны. Важнейшая концепция интеллектуальной Грид-системы - внедрение новейших информационных технологий, цифровой связи, сбора данных, измерений и контроля в энергетические системы для улучшения эффективности их использования. При создании интеллектуальных сетей главная сложность заключается в проектировании сетевой архитектуры, обеспечивающей надёжную и безопасную двустороннюю связь на расстоянии от нескольких метров до нескольких тысяч километров, в том числе до границ других доменов, тоже формируемых с поддержкой «умных» сетей. Двусторонняя связь между удалёнными узлами сетей может поддерживаться различными способами: с использованием проводных слаботочных линий [1, с. 928-997, 1028-1045], проводных линий с передачей данных по силовым кабелям электроснабжения [1, с. 998-1027], а также с использованием беспроводных сетей [1, с. 998-1045]. Телекоммуникационные и информационные технологии играют ключевую роль в интеллектуальных Грид-системах, поскольку необходимо в реальном времени обрабатывать большой объём информации, циркулирующей в пространственно распределённой интеллектуальной системе [20]. Производители промышленного, электрического и коммуникационного оборудования должны работать в кооперации при реализации интеллектуальных Грид-систем. Разработка стандартов для представлений, значений и протоколов транспорта данных является важной составной частью создания этих систем. Конгресс США осознал важную роль стандартов в этой области и в качестве организации, ответственной за координацию и развитие стандартов интероперабельности, назначил NIST – National Institute of Standards and Technology. Технические стандарты для интеллектуальных Грид-систем в настоящее время разрабатывают более 20 американских и международных организаций. Кибербезопасность в интеллектуальной Грид-системе имеет высокий приоритет и должна быть обеспечена уже на архитектурном уровне. Версия 1.0 Руководства по кибербезопасности интеллектуальных Грид-систем была опубликована NIST в августе 2010 г.

Перечислим основные направления и особенности развития современных интеллектуальных Грид-систем на примерах описаний их технических, функциональных и архитектурных особенностей.

В [1, с. 928-951] рассматриваются вопросы интеллектуальной генерации и передачи сигналов с синхронной обработкой данных в режиме реального времени и описываются приложения для системы магистрального электроснабжения (Канада, США), обеспечивающие проведение широкого спектра раз-

нообразных измерений технологических параметров на основе синхронной обработки данных в режиме реального времени с использованием интеллектуальных Грид-систем. В этой же работе определяются основные факторы, влияющие на эффективность измерений, передачи и обработки данных: включение человека или вычислительной системы в контур обратной связи цепи управления, наличие временных задержек передачи и обработки данных, географическая удалённость объектов, пропускная способность каналов передачи и обработки данных, количество обрабатываемой информации, наличие критически важных данных.

В [48, с. 52-61] подчеркнута принципиальная важность решения задачи синхронизации в условиях взаимной зависимости процессов распределения энергии и потоков информации в управляющих интеллектуальных Грид-системах, описана методика планирования полного набора тестов на синхронизм и осуществления измерений в сетях, позволяющая проводить сравнение технологий для реализации синхронизма. Измерения в конкретных сетях для конкретных устройств проводились на основе стандарта IEEE-1588 v2 2008. Показано, что контроллеры PC11588 могут быть взяты за основу управляющих узлов, а LM3S8962 микроконтроллеры на основе ядер ARM Cortex-M3 – в качестве пассивных узлов, чтобы обеспечить требуемый уровень синхронизации в интеллектуальных Грид-системах.

В [1, с. 952-976] представлена многоуровневая парадигма (система понятий) автоматизированного оценивания состояния технических объектов в больших промышленных интеллектуальных сетях. Сети имеют многоуровневую инфраструктуру, обеспечивающую интеграцию подсистем мониторинга и оценки состояния больших взаимосвязанных сетей линий электроснабжения (Испания, США) с использованием «решёток» вычислительных ресурсов. Системы мониторинга функционируют на различных иерархических уровнях, в том числе, на уровнях отдельных электрических подстанций, а также электрических линий передающих сетей. В [1, с. 952-976; 5; 19-20] перечислены типичные трудности, возникающие при организации автоматизированного управления промышленными интеллектуальными сетями: несогласованное разрастание сетей линий электроснабжения в различных регионах и невозможность быстрой перестройки сетевой топологии с синхронизацией взаимосвязей между отдельными частями сетей, приводятся также численные результаты экспериментального оценивания производительности представленной многоуровневой инфраструктуры.

Преобразование данных в знания – главная тема результатов исследований, представленных в [1, с. 977-997]. Оперирование большими массивами данных требует сквозной автоматизации процессов их обработки, начиная от сбора данных до принятия решений в интеллектуальных сетях «решёток» вычислительных ресурсов. Здесь же предлагается новый способ вычислительной обработки данных с ге-

нерацией так называемых переходных (translational) знаний на основе объединения упорядоченных данных с моделями, обеспечивающими эффективность принятия решений, предлагается также решение, реализуемое в системах мониторинга магистральных сетей электроснабжения и при организации управления указанными сетями (штат Техас, США). Кроме того, объясняется процедура преобразования генерируемых знаний в управляющие действия в интеллектуальных сетях электронных устройств. Автоматизированное управление сетями электроснабжения осуществляется по сквозной цепочке: сбор данных – преобразование данных – генерация переходных знаний – управляющие действия.

Роль коммуникации через линии энергоснабжения в технологии интеллектуальных сетей с использованием возможностей Грид-систем исследуется в [1, с. 998-1027]. Сети линий электроснабжения образуют инфраструктуру со значительно более широким и всеобъемлющим использованием компьютерных средств, чем существующая глобальная инфраструктура традиционных проводных и беспроводных коммуникационных сетей. Здесь же анализируется возможность коммуникации через линии энергоснабжения при реализации интенсивно развивающейся технологии поддержки интеллектуальных Грид-сетей («решёток» вычислительных ресурсов) и подчеркивается актуальность поддержки технологии Грид-вычислений в беспроводных сетях [1, с. 998-1045], являющихся наиболее сложным типом сетей с точки зрения обеспечения высокой надёжности и энергетической эффективности коммуникационных процессов. В [1, с. 998-1027] также рассматриваются особенности решения задачи поддержки Грид-вычислений через линии энергоснабжения в распределённых узко- и широкополосных системах с трафиком от нескольких бит в секунду до 200 Мбит/с. В частности, описан метод устранения влияния фединга в коммуникационном канале, учитывающий беспорядочные или периодические изменения мощности сигналов в сети при приёме, позволяющий обеспечить высокую надёжность каналов беспроводной связи.

Обзор потенциальных приложений в интеллектуальных Грид-системах (Smart Grid) и требований к средствам связи в этой инфраструктуре представлен в [48, с. 28-42].

Отмечено, что существующие в мире энергетические сети были спроектированы как централизованные системы, в которых потоки электроэнергии являются однонаправленными от электростанций через линии электропередачи к конечным потребителям. Управление сконцентрировано на крупных узлах распределения энергии, в то время как подстанции и удалённые терминалы почти полностью пассивны. Новые системы распределённой генерации и передачи электроэнергии, объединяющие тысячи генераторов, характеризуемых различными технологиями, уровнями напряжения, тока и мощности, а также с переменной топологией сети, ставят сложные проблемы управления и безопасности в этой изменчивой среде. Технология интеллектуальных Грид-систем

как раз и нацелена на решение этих проблем за счет создания двунаправленной инфраструктуры передачи данных между интеллектуальными компонентами, внедрения сложных информационных и вычислительных методов, новых бизнес-процессов. В этой же работе рассмотрены вопросы стандартизации интеллектуальных Грид-систем и определены организационные, занимающиеся данной деятельностью.

В [20] предлагается новый алгоритм предсказания последовательности действий жителей (пользователей) в «умных» домах за счёт улучшенного распознавания ситуаций (эпизодов). Алгоритм реализован в технологии SPEED (sequence prediction via enhanced episode discovery), в основе которой лежит марковская модель конечного порядка. Подчеркивается, что технология SPEED является методом предсказания не единичного действия, а целой последовательности действий человека. Авторы приводят результаты экспериментального оценивания производительности и эффективности технологии на основе представленного алгоритма, подтверждающие возможность предсказания указанных действий с точностью до 88,3%, что является наиболее высоким результатом по сравнению с результатами использования альтернативных алгоритмов LeZi Update, Active LeZi, C4.5 и IPAM.

В [1, с.1074-1082; 18; 21] утверждается, что применение «микрорешёток» вычислительных ресурсов с «умным» (smart) распределением может упростить реализацию множества функций интеллектуальных Грид-систем, в частности, функций поддержки на требуемом уровне показателей надёжности, самовосстановления и управления загрузкой сети. Разделение распределенных сетей на микрокластеры позволяет значительно уменьшить непроизводительные потери, связанные с неэффективным потреблением электропитания и рассеянием тепла, за счёт отключения части простаивающих или сбойных ресурсов. При описании методики автоматической генерации микрокластеров особое внимание уделяется организации их автономной работы и поддержке эффективного связывания с интеллектуальным распределением вычислительных ресурсов.

По мере развития и распространения концепций «Интернет вещей» (Internet of Things) и «Интеллектуальная пыль» (Smart Dust) [21], возрастает интерес специалистов и разработчиков компьютерных систем к расширению областей использования технологии радиочастотной идентификации (RFID, Radio Frequency Identification). RFID-метка представляет собой микрочип с крошечной антенной для маркировки товара и его радиочастотной идентификации посредством бесконтактного чтения и записи содержащихся в нём данных. Если требуется считывать данные одновременно с нескольких RFID-меток, то необходимо использовать протокол, защищенный от нежелательных конфликтов совместного доступа к их данным. В этой же работе предлагается алгоритм оптимизации процесса параллельного выполнения множества сессий доступа к RFID-меткам, обеспечивающий высокопроизводительный, надёжный и эффективный доступ к их данным в условиях слу-

чайных кратковременных разрывов беспроводной сетевой связи.

Типичными образцами систем на основе «решёток» вычислительных ресурсов являются глобальные сервисы и службы сети Интернет. Однако их общий недостаток (если не предпринимаются особые меры повышения показателей надёжности) – относительно слабая масштабируемость и низкая отказоустойчивость при однократных сбоях в сети.

В [26] предлагается гибридная архитектура IPBGA – основанная на P2P Грид-архитектура, использующая протокол информационного пула «решёток» вычислительных ресурсов и предназначенная для поддержки улучшенного управления ресурсами и эффективного планирования задач. В отличие от обычного непосредственного внедрения технологии P2P-сетей в архитектуру «решёток» новая архитектура основана на применении технологии виртуализации, использовании абстрактных физических ресурсов и выполнении заданий в виртуальном окружении, что обеспечивает построение действительно гибридной системы из P2P-сетей и «решёток» ресурсов с переадресацией информационных запросов из одних наборов ресурсов в другие, в том числе в совместно используемые ресурсы. Управление ресурсами и планирование задач осуществляются с учётом адаптивной настройки динамических, распределенных и гетерогенных параметров сети.

Грид-системы, применяемые в космической области исследований для предсказания погоды [29-30], предназначены для функционирования в пространственном масштабе многих сотен и тысяч километров и отличаются повышенной сложностью алгоритмов управления. В [30] предлагается метод моделирования классов взаимодействующих процессов с использованием стандартизированных нотаций (обозначений) при описании их поведения. Сетевые системы управления, состоящие из классов так называемых активных объектов, поддерживают «внутриклассовое» (принадлежащие одному и тому же классу объекты взаимодействуют между собой), и «межклассовое» (взаимодействуют объекты, относящиеся к разным классам объектов) взаимодействие. Эффективность моделирования подтверждается результатами экспериментального оценивания вычислительной производительности и использования ресурсов памяти контроллера регистрации погодных изменений, разработанного в Национальном агентстве по авионавигации и исследованию космического пространства (NASA, USA).

Особый интерес представляют направления развития Грид-систем, основанные на переносе принципов, методов и структур, традиционно применяемых в социальных организациях, непосредственно в сферу проектирования, функционирования и поддержки распределенных Грид-систем, объединяющих «решётки» ресурсов многих миллионов вычислительных узлов. Структуры с ограниченной универсальностью [8] – это понятие, используемое для описания процесса формирования многоядерных архитектур «облачных» вычислений в программно-аппаратных

платформах распределенных систем. При описании таких структур применяются методы, принципы и практические алгоритмы, действующие в организациях человеческого сообщества, например, для декларирования прав, полномочий, правил управления процедурами идентификации, а также при установлении норм поведения и ответственности за их нарушение. Структуры с ограниченной универсальностью могут применяться при проектировании систем для масштабируемых, высоконадежных и безопасных «облачных» клиентских вычислений.

В [1, с. 1046-1062] предложена трехуровневая модель надежности интеллектуальных Грид-систем, основанная на технологиях, разработанных в авиационной промышленности десятилетием раньше. Модель определяет три главные категории подсистем по степени их влияния на устойчивость системы в целом:

- уровень А (или высший): сбой в таких подсистемах приводит к сбоям среди десятков тысяч узлов Грид,
- уровень В (или средний): отказ в этих подсистемах приводит к распространению сбоев среди сотен или даже тысяч узлов, но в меньшей географической области,
- уровень С (или низкий): сбой в таких подсистемах имеет местное значение.

В основу модели надежности Грид-системы положен принцип создания системы, готовой к реагированию на сбой в любой своей части в любой момент времени, независимо от конкретной причины (технический сбой, ошибка пользователя, вредоносная активность в информационных сетях и т.д.).

Многоэшелонная защита системы управления Грид-сети должна включать брандмауэры, подсистему выявления вторжений, защиту от вирусов, червей, шифрование при хранении и передаче данных. Авторы [1, с. 1046-1062] отмечают необходимость учета особенностей применения в сети нового поколения интеллектуальных датчиков и регуляторов. Это устройства повышенной сложности как на индивидуальном, так и на системном уровне, для которых пока нет документированной истории их обслуживания, поэтому их надежность базируется скорее на расчетах, чем на реальной статистике использования. Уязвимость новых Грид-систем к кибератакам, неадекватным действиям пользователей, превышению нагрузки, потере генерирующей мощности и комбинации этих факторов требует переосмысления базовой архитектуры интеллектуальной Грид-системы.

В работе [49] рассматривается концепция микро-системы Грид как системы распределения низковольтного напряжения, генерируемого множеством распределенных источников энергии (например, микротурбинами, топливными элементами, ветровыми генераторами) с множеством устройств накопления энергии. Выявлены проблемы, связанные с функционированием, мониторингом, управлением и защитой микро-системы Грид. Отмечается, что быстрое развитие технологий связи и информатики позволяет находить эффективные по стоимости реше-

ния по обмену информацией в локальных сегментах и в глобальной сети Грид, управлению доступом и диффузией данных. Эти технологии лежат в основе стратегического подхода поддержки эффективного функционирования интеллектуальных микро-систем Грид в реальном времени. Описанная среда высокого уровня на основе сервисно-ориентированной архитектуры для интегрированного моделирования, мониторинга и контроля микро-систем Грид независима от операционной системы, языка программирования и поставщиков, и поэтому является идеальным кандидатом для эффективной интеграции в существующие системы управления и распределения энергии.

Центральным компонентом метаархитектуры является система MGengine, ответственная за реализацию управления, моделирование и мониторинг функций в географически распределенной системе. Ядром MGengine является куратор процесса, который управляет выполнением сервисов микро-системы Грид и добавляет функциональность, связанную с мониторингом. Традиционно куратор процесса описывается посредством языка бизнес-логики.

Описанная в [50, с. 352-356] архитектура Грид национальной нанотехнологической сети (Грид-ННС) ориентирована на приложения в области нанотехнологий, требующих параллельных вычислений на суперкомпьютерах средней мощности. Поскольку многие университеты и научные центры имеют суперкомпьютеры, объединение ресурсов в единую Грид-систему позволяет многократно увеличить вычислительную мощность, доступную каждой организации-участнику. Особенности реализации проекта и базовые идеи, заложенные в архитектуру позволяют разделять поток управляющих команд, таких как граф задания, от потока данных и исполняемых программ. Сервис распределения нагрузки, который распределяет задачи между компьютерами, не управляет потоком данных. Все данные непосредственно передаются со специальных серверов хранения на компьютерные ресурсы и обратно. Другим важным отличием от традиционных решений является использование архитектурного подхода REST для проектирования сервисов. Например, сервис распределения нагрузки RESTful-грид-сервис позволяет упростить протоколы обмена между центральными сервисами по сравнению с традиционными решениями.

Подробную информацию о функционировании Грид-ННС, а также инструкции для пользователей и администраторов Грид-ННС можно получить на сайте проекта <http://ngrid.ru>.

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ И АЛГОРИТМЫ ДЛЯ ГРИД-СИСТЕМ

Показатели эффективности Грид-систем определяются не только концептуальными, архитектурными и функциональными характеристиками аппаратных средств, но и в значительной мере свойствами программного обеспечения. Перечислим некоторые направления развития программного обеспечения интеллектуальных Грид-систем в типичных примерах описаний их моделей, распределенных сервисов, методов обработки запросов.

В качестве примера модели программной инфраструктуры, разработанной на языке Java и предназначенной для поддержки «решёток» вычислительных ресурсов, может рассматриваться модель, представленная в [22]. Модель включает систему понятий и программный инструментарий для поддержки распределенных «решёток» вычислительных ресурсов и может использоваться при разработке ПО промежуточного уровня и поддерживает совместное использование ресурсов очень большого количества ПК в распределенной сети, в том числе в сети Интернет с подключением к «решётке» одновременно миллионов пользователей.

Реализация параллельного распределения файлов в иерархии программных агентов описывается в [23]. Распределенные сети в виде «решёток» вычислительных ресурсов привлекательны с точки зрения возможности практически неограниченного увеличения количества подключенных к сети ПК, использующих параллельное ПО. Проблемы возникают при распределенной передаче множества необходимых файлов на удалённые ПК, на которых может оказаться недостаточно свободного места на носителях данных, или носители могут быть временно отключены от сетевой файловой системы, или сетевые устройства могут оказаться неточно синхронизированными по времени включения. В этой работе предлагается ПО промежуточного уровня AgentTeamwork, предназначенное для установки на удалённые ПК набора иерархически организованных программных агентов, обеспечивающих запуск, мониторинг, перезапуск и контроль выполнения параллельных процессов с регистрацией прохождения их контрольных точек.

Пример применения дерева объединения сигнатур для вычислительной обработки объектно-ориентированных запросов представлен в [24]. Опыт применения сигнатурного метода поиска с полнотекстовым индексированием при вычислительной обработке объектно-ориентированных запросов описывается для поиска в очень больших массивах данных. В качестве индексов используются файлы сигнатур, формирование и кодовое представление которых осуществляется посредством динамической балансировки дерева типа B+, получившего название дерева объединения (декластеризации) сигнатур. Все сигнатуры, соответствующие запросу, распределенные по «ветвям» и «листьям» дерева, при обработке запросов могут быть кумулятивно найдены, отобраны и затем объединены (декластеризованы), начиная с одного узла и охватывая всю область поиска в множестве больших объектно-ориентированных БД.

Основная проблема при организации сети с поддержкой «решёток» узлов состоит в достижении необходимых показателей эффективности и самоорганизации системы управления распределенными вычислительными ресурсами. Требуется решение также сложная задача управления масштабируемыми, гетерогенными сетями и сетями с непредсказуемым поведением, например, с периодическим отключением части вычислительных ресурсов при разрыве беспроводных сетей. В [26] предлагается метод фор-

мального описания и специфицирования указанного процесса управления посредством использования готовых шаблонов выбора групп множества автономных программных агентов, обеспечивающих эффективное распределенное управление сгруппированными подсистемами «решёток» узлов.

Эффективное планирование задач параллельной обработки данных в потоках заданий также является одним из условий эффективной работы Грид-систем. Использование моделей потоков заданий позволяет осуществлять автоматическое и прозрачное для пользователей планирование выделения ресурсов для выполняемых вычислительных задач, в том числе с гарантированной поддержкой требуемого уровня качества услуг. В [27] представлен метод планирования задач параллельной обработки данных в потоках заданий на основе оценок временных и стоимостных затрат процессов, включенных в обработку. Показано, что алгоритмы обеспечения доступа к вычислительным ресурсам могут использоваться в распределенных вычислительных сетях (ВС) с поддержкой связи через Интернет.

При решении задач в интеллектуальных распределенных «решётках» вычислительных ресурсов процессы вычислительных заданий должны адаптироваться к динамически изменяющемуся гетерогенному окружению. Адаптивное выполнение заданий в окружении «решёток» вычислительных ресурсов предоставляет дополнительные возможности для повышения общей эффективности Грид-систем. Основным критерием эффективности адаптации – обеспечение качества услуг (сервисов), предоставляемых пользователям ВС. В [28] предлагается интегрированная инфраструктура (система понятий, алгоритмы и ПО) для управления ресурсами в распределенных «решётках» с поддержкой многоагентной системы, распределяющей вычислительные задания по узлам сети. В случае перегрузки или «зависания» какого-либо узла сети, а также при общем снижении показателей качества сервисов, система перераспределяет вычислительную нагрузку таким образом, чтобы показатели качества сервисов непрерывно сохранялись на требуемом уровне.

Поставщиков различных «облачных» услуг, особенно базовых приложений и инфраструктурных сервисов, существует много, однако конкретные предложения даже самых известных из них [51-55] сильно различаются. Даже в отношении того, каким образом рассчитывать стоимость услуг в «облаке», единого мнения пока нет. Однако, как правило, клиенты могут выбирать интересующие их параметры настроек соединений (в том числе, уровни безопасности, вид управления доступом и используемые API-функции) и оплачивать услуги в соответствии с объемом потребляемых ресурсов.

В сервисах Amazon [51] плата за самый известный инфраструктурный сервис Elastic Compute Cloud взимается в зависимости от количества потребляемого трафика. Услуги электронной почты Google Gmail [52] для личного пользования предоставляются бесплатно. Для корпораций существуют две версии Google Apps: бесплатная базовая (имеет ограничен-

ный функционал) и платная профессиональная (с расширенными возможностями). Технология Google App Engine позволяет создавать на серверах Google виртуальные Web-приложения на языках Java или Python. Менее известный сервис Skytap от Skytap Virtual Lab [53] предоставляет платформу, на которой клиенты могут запускать виртуальные машины и приложения, не создавая виртуальную инфраструктуру самостоятельно.

Компания VMware [54], занимающая лидирующие позиции на рынке технологий виртуализации, предлагает продукт vSphere, предназначенный для создания частного пользовательского «облака». Компании VMware с поддержкой целого ряда влиятельных партнеров удалось выпустить продукт, который может помочь клиентам в создании среды, объединяющей как частные, так и общедоступные «облака».

Программный продукт Microsoft Azure [55], представляющий собой гипервизор, построенный на платформе Windows Server 2008, конкурирует с программным обеспечением виртуализации VMware, а платформа Azure предлагает комплексный способ разработки программ для выхода в «облако» сервисов и ресурсов хранения данных с использованием СУБД SQL Azure Web Edition, включая поддержку сообщений Service Bus и маркеров Access Control. В подходе, предложенном в [50, с. 114-120], используется программное обеспечение промежуточного уровня – система g-Lite [56], которая создана и реализована в архитектуре EGEE [57]. Программное обеспечение поддерживает Грид-сервисы брокерингом ресурсов, вычислением заданий и хранением данных. Службы безопасности имеют механизм для идентификации, авторизации, верификации пользователя или сервиса и ведут журналы для анализа и аудита. Сервисы мониторинга содержат механизм для мониторинга задач, определения необходимых ресурсов и вывода служебной информации. Сервисы управления нагрузкой включают вычислительные ресурсы, на которые размещаются задания для выполнения, планировщик задач, и учетную информацию. Сервисы данных имеют модуль хранения для доступа к ресурсам хранения данных и службу передачи файлов. ПО промежуточного уровня обеспечивает сервисы для сбора учетной информации. Эта информация содержит данные о количестве размещенных заданий, пользователей, разместивших задания, и о виртуальных организациях, к которым принадлежат эти пользователи.

ПО g-Lite не основано на сервисно-ориентированном подходе, который требует реализации таких архитектурных принципов, как контракт, абстракция, повторная используемость, способность к композиции и определимости служб. Под контрактом понимается документ, описывающий, каким образом сервис может быть доступен с помощью программного модуля. Требования абстракции обязывают, чтобы сервисы отражали только логику, определенную в сервисном контракте, и скрывали ее реализацию от клиента. Повторная используемость предполагает использова-

ние сервисов многократно и многими пользователями. Композиция обозначает возможность группирования сервисов в сложные сервисы и выполнение их как одного процесса. Определимость обеспечивается стандартным механизмом для определения сервисов.

ПО g-Lite не отвечает данным требованиям. Например, службы безопасности в системе g-Lite не имеют стандартного описания, как и механизмов определения сервисов. Однако ее создатели и не стремились к этому. Целью их проекта было создание и реализация Грид-архитектуры для научных исследований. Проект имел два приоритетных направления: обслуживание биомедицинских экспериментов и экспериментов в области физики высоких энергий. Последняя область связана со сложными бизнес-процессами. Бизнес-процессом считается множество сервисов, упорядоченных в общей схеме их выполнения. Процессы в физике высоких энергий включают не только службы из Грид-инфраструктуры, но и компоненты специального ПО для моделирования и анализа данных.

В [58] рассматриваются вопросы оптимизации потребления энергии в крупномасштабных вычислительных системах, таких как распределенные системы Грид, и предложено решение на основе кооперативной теории игр. Детально исследуется частная задача распределения работ с учетом потребления энергии на вычислительных узлах в Грид-системе, каждый из которых оборудован устройством динамического управления напряжением. Математически задача формулируется как многоцелевое с множеством ограничений расширение обобщенной задачи о назначениях. Решение проблемы основано на решении задачи Нэша о переговорах. Фактически, это первая исследовательская работа, которая рассматривает проблему оптимизации управления ресурсами в Грид-системе с позиции теории игр. Имитационное моделирование показывает, что предложенная методика имеет большую эффективность, чем ранее предложенные подходы на основе «жадных» эвристик или эвристик с линейной релаксацией.

В [59] теория игр рассматривается как ключевое аналитическое средство проектирования интеллектуальных Грид-систем. Теория игр является формальной аналитической и концептуальной основой с множеством математических методов, дающих возможность изучать сложные взаимодействия между независимыми рациональными игроками. Детально рассмотрено применение этой теории в конструировании микросистем Грид, систем управления спросом в энергетических сетях, а также обеспечении эффективной передачи данных для реализации интеллектуального управления. Большая часть работ в этой области основывается на классических, статичных некооперативных играх. В работе проводится мотивация перехода к исследованиям на основе динамических игровых моделей, как кооперативных, так и некооперативных. Интерес представляют также байесовы игры, в которых игроки имеют ограниченную информацию о целевых функциях и стратегиях

своих оппонентов. Это соответствует специфике интеллектуальных Грид-систем, когда трудно оценить стратегии и цели других игроков. Теория игр может быть использована для анализа уязвимостей интеллектуальных Грид-систем на уровнях инфраструктуры, связи, передачи пакетов данных и оценки текущего состояния системы. Для этого может быть полезен формализм дифференциальных игр.

В [60] рассматривается применение технологии активного предупреждения о сбоях AFA (active-fault-alarm) к положительной оптической сети на основе 10 гигабитной сети Ethernet (10G-EPON). Эта система широко используется для реализации интеллектуальных Грид в области сетей распределения электрической энергии. Разработана модификация AFA-PP в виде схемы заранее предусмотренной защиты от сбоев, позволяющая более эффективно восстанавливать сеть 10G-EPON при реальных отказах, а также распознавать ситуации с сигналами ложной тревоги, поступающими от сенсорной сети контроля и мониторинга. Доля ложных сигналов может достигать 45% при количестве датчиков около 100 тыс. единиц. Разработанные алгоритмы с применением методов искусственного интеллекта позволяют существенно повысить эффективность всей системы управления распределением электроэнергии за счет решения проблемы ложной тревоги.

В [61] проводится сравнительный анализ двух главных методик для обеспечения поддержки интенсивных параллельных приложений в глобальной Грид-системе. Эти методики позволяют сгладить скачкообразный характер процессов в сетях Грид, обусловленный непредсказуемостью и динамическим характером доступных процессорных и сетевых ресурсов. Сравнение осуществлялось на основе имитационного моделирования. Показано, что существует простая для измерения статистика Y и пороговое значение Y_1 , такие, что метод Динамического Балансирования Нагрузки (ДБН) превосходит метод Репликации Работ (PP), когда $Y > Y_1$, и, наоборот, второй метод превосходит первый при $Y < Y_1$. На основе данного наблюдения авторы предложили простой в реализации подход, который динамически выбирает метод управления на основе контроля порогового значения статистики. Показано, что гибридный метод лучше справляется с непредсказуемостью нагрузки в глобальной Грид-системе, чем методы ДБН или PP по отдельности.

В [62] описана проблема, связанная с тем, что существуют эгоистичные или враждебные узлы Грид-системы, которые потребляют ценные ресурсы сети, но не вносят необходимого вклада в разделение нагрузки сети, либо пытаются разрушить сеть. Это приводит к сильной деградации эффективности сети и ограничивает ее развитие. Для того чтобы противостоять данным угрозам, разработана среда адаптивного управления ресурсами Грид – QGrid, которая интегрирует фактор доверия в управляемый экономическими методами процесс распределения ресурсов. Каждый провайдер распределяет ресурсы со-

гласно аукционной цене и значению доверия запрашивающей ресурс организации, контролируя соответствующие пороговые значения. Неполнота информации является главной проблемой при определении пороговых значений. Для ее решения предлагается использовать метод Q-обучения, который способен адаптировать динамику Грид-структуры. Вводится также простая изоляционная схема, ограничивающая возможности враждебных узлов к присоединению к системе Грид. Прототип QGrid был успешно реализован в реальной архитектуре CROWN, предназначенной для тестирования разработанных Грид-систем. Теоретический анализ и обширные эксперименты продемонстрировали эффективность QGrid.

В [4] предложена система AID (Automated Incident Detection), которая усиливает автотранспортные спонтанные сети VANETs и реализована как приложение VGrid. В пределах некоторой географической области спонтанная сеть транспортных средств, оборудованная сенсорами и компьютерами, может образовывать кластер сенсоров и компьютеров, составляющих вместе VGrid – инфраструктуру контроля и управления, которая собирает и обрабатывает данные трафика в реальном времени для принятия управленческих решений. Показано, что такая распределенная система сбора и обработки данных существенно улучшает возможность фиксировать в реальном масштабе времени транспортные происшествия на дорогах. При низкой цене устройств система может быть также экономически эффективной.

В [63] представлен алгоритм Robber, предназначенный для максимизации пропускной способности между множеством кластеров компьютеров в Грид-сети. Этот подход основан на идеях протокола BitTorrent. Узлы в одном кластере компьютеров работают коллективно, пытаясь захватить данные с равноправных кластеров. Алгоритм Robber автоматически адаптируется к полосе пропускания сети. В пределах кластера узлы автоматически подстраиваются к количеству данных, которые они могут захватить дистанционно согласно их относительной эффективности. Показано, что данный алгоритм обеспечивает доставку всех данных всем узлам сети. Алгоритм Robber реализован в основанной на языке Java среде Ibis.

В [5] утверждается, что компании США, поддерживающие городские сети электроснабжения (ГСЭ) могут получать дополнительные прибыли и эксплуатационные возможности на основе применения программного обеспечения преобразования исторически накапливаемых данных о состоянии ГСЭ в модели предсказания рисков отказов компонентов и систем. Указанные модели могут непосредственно использоваться компаниями, поддерживающими ГСЭ, в частности, при расчёте показателей надёжности сетей и определении приоритетов для выполнения эксплуатационных и ремонтных работ. Здесь же описывается накопленный практический опыт использования программных моделей в системе ГСЭ Нью-Йорка (США).

ЗАКЛЮЧЕНИЕ

Методология Грид – одно из перспективных направлений развития современных технологий разработки и использования программного обеспечения в распределенной компьютерной среде с принципиально новой организацией вычислений и управления потоками заданий и данных. Представленный обзор ряда направлений исследований в области Грид-систем подтверждает возрастание интереса к этой технологии. Сегодня приложения инфраструктуры Грид, в частности «облачные» сервисы, позволяют интегрировать очень большой объем географически удаленных компьютерных средств и предоставлять пользователям мощные, разнообразные, никогда ранее не доступные, удобные и относительно недорогие вычислительные и коммуникационные ресурсы.

Перечисленный в обзоре ряд решений в области разработки и внедрения технологии Грид является показателем современного мирового уровня и тенденций развития технологических инноваций в рассматриваемой сфере. Сегодня для поддержки и развития технологии Грид существуют разнообразные инструментальные средства, доступные широкому кругу разработчиков программ и пользователей средств вычислительной техники, в том числе в Интернете. Крупные транснациональные корпорации вкладывают огромные материальные и интеллектуальные ресурсы, оказывающие реальное влияние на производительность труда, стиль коммуникации и образ мышления многих миллионов пользователей и разработчиков программного обеспечения.

СПИСОК ЛИТЕРАТУРЫ

1. Proc. IEEE. – 2011. – Vol. 99, №6. – P. 928-951, 952-956, 977-997, 998-1027, 1028-1045, 1074-1082, 1098-1115, 1046-1062.
2. Кафедра информатики, Факультет нанотехнологий и информатики, МФТИ. - 2007 г. – URL: <http://www.fnti.kiae.ru/>.
3. NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 1.0, Office of the National Coordinator for Smart Grid Interoperability. – URL: http://www.nist.gov/public_affairs/releases/upload/smartgrid_interoperability_final.pdf.
4. Khorashadi B. et al. Distributed Automated Incident Detection with Vgrid //IEEE Wireless Communications -2011. – Vol.18, №1. – P.64-73.
5. Rudin C., Waltz D., Anderson R. et al. Machine Learning for the New York City Power Grid // IEEE Trans. Pattern Anal. and Mach. Intell. – 2012. – Vol.34, №1-3. – P. 328-345.
6. Chang R., Yu Yu, Hairong Lv. et al. Selling the Smart Grid - Part 1 // IEEE Consum. Electron. Mag. – 2012. – Vol.1, №2. – P. 24-31.
7. Пройдаков Э.М., Теплицкий Л.А. Англо-русский толковый словарь терминов и сокращений по ВТ, Интернету и программированию. Computers (En-Ru) (к версии ABBYY Lingvo x3). 5-е изд., испр. и дополн. – М., 1998-2007.
8. Hewitt C. ORGs for Scalable, Robust, Privacy-Friendly Client Cloud Computing // IEEE Internet Comput. – 2008. – Vol.12, №5. – P. 76-99.
9. Каляев А.И. Об одном способе решения поступающего потока связанных задач в мультиагентных GRID //Проблемы управления и моделирования в сложных системах: сб. трудов. – Таганрог, 2013. – С.409-413.
10. Фогарти К. Облачные вычисления: определения и решения // Журнал «Директор информационной службы». – 2011, №3. ЗАО "Издательство «Открытые системы». – URL: <http://www.osp.ru/cio/2011/03/13007508/>.
11. SETI@home (Search for ExtraTerrestrial Intelligence at home). – URL: <http://setiathome.berkeley.edu/>.
12. Folding@home distributed computing project. – URL: <http://folding.stanford.edu/>.
13. DIMES – a distributed scientific research project. – URL: <http://www.netdimes.org/new/>.
14. The Clean Energy project. – URL: <http://cleanenergy.harvard.edu/>.
15. Hydrogen@Home. – URL: <http://hydrogenathome.org/>.
16. World Community Grid. – URL: <http://www.worldcommunitygrid.org/>.
17. PlanetQuest. – URL: <http://www.planetquest.org/>
18. Zhabelova G., Vyatkin V. Multiagent Smart Grid Automation Architecture Based on IEC 61850/61499 Intelligent Logical Nodes // IEEE Trans. Ind. Electron. – 2012. – Vol.59, №5. – P. 2351-2362.
19. Mo Y., T. Kim Hyun-Jin, Brancik K. et al. Cyber-Physical Security of a Smart Grid Infrastructure // Proc. IEEE. – 2012. – Vol.100, №1. – P. 195-209.
20. Alam Muhammad Raisul, Reaz I., Mohd Ali. SPEED: An Inhabitant Activity Prediction Algorithm for Smart Homes // IEEE Trans. Syst., Man. and Cybern. A. – 2012. – Vol.42, №4. – P. 985-990.
21. Fyhn Karsten, Jacobsen Rasmus Melchior, Popovski Petar, Torben Larsen. Fast Capture—Recapture Approach for Mitigating the Problem of Missing RFID Tags // IEEE Trans. Mobile Comput. – 2012. – Vol.11, №3. – P. 518-528.
22. Niculiță R., Țăpuș N. Model for a Grid Infrastructure Based on Java // Sci. Bull. C / Politehn. Univ. Bucharest. - 2009. – Vol.71, №4. – P. 113-124.
23. Fukuda Munehiro, Miyauchi Jumpei. An implementation of parallel file distribution in an agent hierarchy // J. Supercomput. – 2009. – Vol.47, №3. – P. 255-285.

24. Nadarajan R. Applying SD-Tree for Object-Oriented Query Processing // *Informatica (Slovenia)*. – 2009. – Vol.33, №2. – P. 169-179.
25. Li D. et al. IPBGA: a hybrid P2P based grid architecture by using information pool protocol // *J. Supercomput.* – 2009. – Vol.49, №2. – P. 159-189.
26. Chao I., Ardáiz Ó., Sangüesa R. A Group Selection Pattern Applied to Grid Resource Management // *Knowl. Eng. Rev.* – 2009. – Vol.3, №1. – P. 91-103.
27. Ranaldo N., Zimeo E. Time and Cost-Driven Scheduling of Data Parallel Tasks in Grid Workflows // *IEEE Syst. J.* – 2009. – Vol.3, №1. – P. 104-120.
28. Roy S. Mukherjee N. Adaptive Execution of Jobs in Computational Grid Environment // *J. Comput. Sci. and Technol.* – 2009. – Vol. 24, №5. – P. 925-938.
29. Shinohara M., Kikuchi T. Space Weather News on the Web at NICT // *J. Nat Inst. Inf. and Commun. Technol.* – 2005. – Vol.52, № 3-4. – P. 207-214.
30. Goel Ankit et al. Interacting process classes // *ACM Trans. Software Eng. and Methodol.* – 2009. – Vol.18, №4. – P. 13/1-13/47.
31. Burns Catherine M., Bryant David J., Chalmers Bruce A. Boundary, Purpose, and Values in Work-Domain Models: Models of Naval Command and Control // *IEEE Transactions on Systems, Man, and Cybernetics - Part A, Systems and Humans.* – 2005. – Vol. 35, №5. – P. 603-617.
32. Ke Hongfa, Chen Yongguang, Wang Guoyil. Target Assignment Model of Electronic Equipment Based on GM(1,1) Model // *The Journal of Grey System 3 – National University of Defense Technology; Changsha, China, 2005.* – P. 235-242.
33. Liu X. L., Liao G. S. Multi-target localisation in bistatic MIMO radar // *Electronics Letters.* – 2010. – Vol. 46, №13. – P. 945-949.
34. IBM – URL: <http://www.ibm.com/>.
35. Samsung Electronics – URL: <http://www.samsung.com/>.
36. Apple – URL: <http://www.apple.com/>.
37. SAP – URL: <http://www.sap.com/>.
38. Microsoft Corp. – URL: <http://www.microsoft.com/>.
39. DIMES – a distributed scientific research project. – URL: <http://www.netdimes.org/new/>.
40. The Clean Energy project. – URL: <http://cleanenergy.harvard.edu/>.
41. Galli S., Logvinov O. Recent development in the standardization of powerline communications within the IEEE // *IEEE Commun. Mag.* – 2008. – Vol. 46, № 7. – P. 64-71.
42. Gharavi H., Ban K. Multihop sensor network design for wideband communications // *Proc. IEEE.* – 2003. – Vol. 91, № 8. – P. 1221-1234.
43. Royer E., Perkins C. Multicast operation of the ad-hoc on-demand distance vector routing protocol // *Proc. ACM Mobicom'99, Seattle, WA.* – Aug. 1999. – P. 207-218.
44. Sinha P., Sivakumar R., Bharghavan V. MCEDAR: Multicast core extraction distributed ad-hoc routing // *Proc. Wireless Commun. Netw. Conf., New Orleans, LA.* – 1999. – Vol. 3. – P. 1313-1317.
45. Zheng J., Lee M. J. A comprehensive performance study of IEEE802.15.4 // *Sensor Network Operations, Hoboken, ch. 4.* – 2006. – P. 218-237.
46. Motegi J., Horiuchi H. AODV-based multipath routing protocol for mobile ad hoc networks // *IEICE Trans. Commun.* – 2004. – Vol. E87-B, №9. – P.2477-2483.
47. Tsirigos and Haas Z. J. Multipath routing in the presence of frequent topological changes // *IEEE Commun. Mag.* – 2001. – Vol. 39, №. 11. – P. 132-138.
48. *IEEE Trans. Ind. Inf.* – 2013. – Vol.9, № 1. – P. 28-42, 52-61.
49. Vaccaro A. et al. An Integrated framework for Smart Microgrids modeling, monitoring, control, communication, and verification // *Proc. IEEE.* – 2011. – №1. – P.119-132.
50. Демичев А.П. и др. Архитектура Грид для национальной нанотехнологической сети // *Distributed Computing and Grid-technologies in Science and Education. Proceedings of the 4 International Conference, Dubna, July 3.* – 2010. – P. 352-356.
51. Amazon.com, Inc. – URL: <http://www.amazon.com/>.
52. Google. – URL: <http://www.google.ru/>.
53. Skytap. – URL: <http://www.skytap.com/skytap-cloud/>.
54. VMWare, Inc. – URL: <http://www.vmware.com/ru/>.
55. Microsoft Corp. – URL: <http://windows.microsoft.com/ru-RU/>.
56. Programming the Grid with g-Lite. – URL: <http://cgsweb.cern.ch/record/936685/files/egee-tr-2006-001.pdf>
57. EGEE. – URL: <http://public.eu-egee.org>.
58. Khan S., Ahmad I. A cooperative Game Theoretical Technique for Joint Optimization of Energy Consumption and response time in Computational Grids // *IEEE Transaction on Parallel and Distributed Systems.* – 2009. – Vol.20, №3. – P.346-360.
59. Walid Saad, Zhu Han, H. Vincent Poor, Tamer Basar. Game-Theoretic Methods for the Smart Grid // *IEEE Signal Processing Magazine.* – 2012. – Vol.29, № 5. – P. 86-105

60. Bai H.F., Wang B.T., Wang D S., Li M.W. Active-fault-alarm enabled pre-protection scheme in 10G-EPON for smart power grid // J. China Univer. Posts and Telecommun. – 2012. – Vol.19, №4. – P. 106-109. IEEE Transaction on Parallel and distributed Systems. – 2009. – Vol. 22, №2. – P.231-244.
Материал поступил в редакцию 30.09.13.
61. Dobber M. et al. Dynamic Load balancing and job replication in a Global-scale grid Environment: A comparison // IEEE Transaction on Parallel and distributed Systems. – 2009. – Vol.20, №2. – P.207-218.
62. Li Lin, Jinpeng Huai. QGrid: An Adaptive Trust Aware Resource Management Framework // IEEE Systems Journal. – 2009. – Vol.3, №1. – P.78-90.
63. Burger M. den, Kielmann T. Collective receiver-initiated multicast for Grid Applications //

Сведения об авторах

ГОЛУБЕВ Владимир Михайлович – старший научный сотрудник ВИНТИ РАН, Москва
e-mail: are@viniti.ru

ДУДИН Евгений Борисович – кандидат технических наук, зав. Отделом ВИНТИ РАН
e-mail: dudine@viniti.ru

УШАКОВ Валерий Николаевич – кандидат технических наук, старший научный сотрудник ВИНТИ РАН
e-mail: are@viniti.ru

Д.А. Кузьмина

О пространственной привязке документов в области рудной геологии

Рассмотрены вопросы привязки географических и геолого-географических данных в русскоязычных текстах. В обобщенном виде представлены состав, специфика и возможные неоднозначности при автоматизированном выделении географической лексики из текстов документов в области рудной геологии с учетом ее пространственного определения на координатной основе. Кратко показаны общая структура и основные принципы построения геоинформационной системы для анализа пространственной привязки.

Ключевые слова: геоинформационная система, анализ текста, географическая лексика, автоматическое индексирование, геология рудных месторождений, база данных, электронные библиотеки

ВВЕДЕНИЕ

Специфика информации в геологических дисциплинах, предметных областях науки о Земле, а также в исследованиях экологической направленности, заключается в её географической привязке к объекту изучения. Количество пространственно привязанной геологической информации составляет, как минимум, более половины массива геоданных. Территориальный аспект в геологии несет основную нагрузку и концентрирует наибольший исследовательский интерес пользователей электронных библиотек. Ведение пространственного аспекта геоданных регламентируется только для картографических данных и спутниковых снимков. Значительные объемы текстов – научных публикаций, монографий, отчетов и других видов документов – индексируются рубрикой классификатора, зачастую укрупненной и не всегда обязательной для индекса документа. Извлечение из текста географического компонента, как и любой другой полезной информации, – инструмент для последующего выполнения аналитических процедур, обеспечивающих, в первую очередь, поиск информации пользователем. Поиск документов по географической привязке в электронных хранилищах данных наиболее целесообразно проводить с использованием геоинформационных систем (ГИС), позволяющих устранять территориальную вложенность, а также проводить комплексный анализ сведений. Внедрение ГИС-сервисов концентрирует интерес отечественных исследователей к вопросам пространственной составляющей данных и диктует повышенные требования к автоматизированному выделению географической привязки из текста, ее корректности и однозначности при переходе на картографическую основу.

Среди проблем, связанных с терминологическим обеспечением баз данных, содержащих сведения о

геологических исследованиях, следует отметить необходимость разработки системы подходов к индексированию пространственной информации, несущей полидисциплинарную нагрузку, а также учета тематико-географической устойчивой лексики (имена собственные геообъектов, районирование в узкоспециализированных предметных областях и т.п.) и географической привязки территорий, не имеющих четких границ.

Задача пространственного определения текстов по геологическим изысканиям базируется на ведении кадастра географических наименований, оперировании стандартами терминологии и картографии, построении онтологий терминов, ведении базы ретроспективных пространственно-лексических данных, а также на дисциплинарных знаниях и оперирует вероятностью приуроченности исследования к территории.

Рассмотрим подробнее вопросы автоматизированного выделения пространственной привязки из массива документографических данных для последующего анализа с использованием ГИС-технологий, а также проблемы и возможные неоднозначности, связанные с геологической спецификой текстов.

ОБЩИЕ ПОДХОДЫ К ИНДЕКСИРОВАНИЮ ПРОСТРАНСТВЕННОЙ ПРИВЯЗКИ ТЕКСТОВЫХ ДАННЫХ

В настоящее время в практику обслуживания информационных массивов по геологии внедряется использование пространственного геокодирования с заданием координат и выбором областей (полигонов) поиска информации, что технологически расширяет возможности полезной отдачи геологических данных. Известны реализации использования ГИС для хранилищ данных. (Например, в геологии горючих полезных ископаемых – <http://sis.slb.ru/sis/item398/> или отечественные разработки для поиска публикаций, например, в области природопользования «Природные ресурсы Карелии» – <http://ias.krc.karelia.ru>) [1,

с.121-128, 145-151]). Для эффективного использования ГИС необходимо обобщение пространственной составляющей текстовых данных на основе унифицированной модели, учитывающей тождества и онтологические связи между территориальными сущностями различных документов с устранением неоднозначностей при переходе на координатную основу, обеспечивающую получение характеристик информационного наполнения хранилища данных. Эффективность использования ГИС обусловлена рядом взаимосвязанных операций над массивом документов, которые условно можно сгруппировать по следующим направлениям:

- **извлечение** географической составляющей из текстов описаний исследований;
- **установление онтологий**, связей и тождеств между пространственной терминологией и пространственными сущностями различных документов;
- **картографическое отображение**, в том числе с учетом определения территорий, указанных в неявной форме.

В информационном хранилище документы индексируются по наиболее семантически значимой информации (названию, реферату), отраженной в базе метаданных наряду с данными об авторах и другими атрибутами. В наиболее распространенном варианте указание пространственной привязки документа традиционно опирается на кадастры географических наименований. В общем случае таковыми являются Общероссийский классификатор стран мира (ОКСМ) и Общероссийский классификатор объектов административно-территориального деления (ОКАТО). В фондах научной литературы географическая привязка индексируется по укрупненным рубрикам с использованием таких общепринятых классификаций, как Библиотечно-библиографическая классификация (ББК), Универсальная десятичная классификация (УДК), Государственный рубрикатор научно-технической информации (ГРНТИ) и Рубрикатор ВИНТИ по разделу «География». К достоинствам структуры и методики построения индекса рубрик УДК можно отнести возможности отражения вложенности континентов, стран, их субъектов, онтологии природных вод (притоки рек, деление мирового океана, кодовое отражение принадлежности к северным и южным морям), а также указания физико-географического региона. Однако даже при ответственном и корректном кодировании документа поиск в массиве и тем более ГИС-реализация в электронном каталоге могут происходить по укрупненным составляющим (например, экологические исследования по регионам и областям России в [2]). Такой подход приемлем, возможно, для пространственных характеристик проектных исследований масштабного комплексного характера по региональным, областным проблемам, содержащих общераспространенную географическую лексику. Использование УДК недостаточно при обработке текстов по исследованиям геологических объектов, в большинстве своем не отраженных в данной классификации. Заметим, что на практике форматы метаданных электронных библиотек не всегда регламентируют обязательное отображение пространственной составляющей. При добавлении геоинформационного сервиса

необходим перевод кодового или текстового обозначения в географические координаты, который может основываться на сопоставимости форматов различных массивов [3] для гетерогенных данных, и чаще предполагает повторную обработку текстов с целью извлечения пространственной составляющей.

Однозначная территориальная привязка реализуется в географических координатах. Если фактографические пространственные данные, космоснимки, карты фиксируют координатную составляющую, то критерием пространственной привязки текстов документов является географическая лексика в названии, тексте рефератов публикаций, преобразованная в числовой или текстовый (ключевое слово или словосочетание) индекс документа. Как и в общем случае автоматизированного анализа текста, извлечение пространственной привязки из названий и рефератов статей, монографий, текстов и баз данных и др. информационных ресурсов, зависит от сложности языка, качества текста с точки зрения грамотности изложения, стилевой и жанровой однородности, а также априорных предметных знаний [4].

Автоматизированная обработка пространственных данных является элементом анализа текста, в качестве простого варианта которого можно обозначить выделение ключевых слов и словосочетаний на основе специализированных словарей и тезаурусов. Зарубежные технологии выделения пространственной составляющей преимущественно основаны на использовании регулярных выражений, но недостаточно эффективны для русскоязычных текстов. Это в значительной мере обусловлено грамматическими правилами построения предложений, словосочетаний, а также контекстной, тематической зависимостью географических терминов в текстах на русском языке. Тематическая классификация играет немаловажную роль при анализе пространственной лексики. Подходы к оценке и составу тематической классификации зависят от объема массива базы данных и его предметной специализации. В текстовом анализе, построенном на обучаемых методах, графематические (например, первая заглавная), морфологические (часть речи, род, число, падеж) признаки, критерии принадлежности к словарям (по тематической категории) и контекст, могут использоваться в статистическом анализе новой терминологии [1, с. 37-42; 254-258]. Применяется также вероятностное определение принадлежности текстового фрагмента, содержащего географическую лексику, к определенной территории на карте. Наиболее перспективным представляется подход, рассматривающий атрибуты и полидисциплинарный контекст терминов, описанный в [4].

Ввиду требований позиционирования на фоне грамматических правил существующих неоднозначностей терминов географической лексики между собой и с терминами естественного языка, автоматизированное выделение географической составляющей из текстов не может быть основано только на графематических и морфометрических признаках со статистической обработкой и использованием словарей и кадастров географической лексики и должно базироваться, как минимум, на тезаурусе, учитывающем онтологию территориальных, дисциплинарных и общенаучных

терминов. Целесообразность применения такого лингвистического обеспечения рассмотрим на примерах неоднозначностей, возникающих при автоматизированном извлечении пространственной привязки из названий и текстов документов в по рудной геологии.

ГЕОГРАФИЧЕСКАЯ ЛЕКСИКА В ДОКУМЕНТАХ ПО ГЕОЛОГИИ РУДНЫХ МЕСТОРОЖДЕНИЙ

Пространственная привязка документов по геологии рудных месторождений представлена географической лексикой, обозначающей территориальные и исследовательские объекты различного масштаба, преимущественно отражающей предметную (рудные месторождения), физико-географическую, а также административно-региональную составляющую. Названия и описания информационных ресурсов оперируют общепринятой географической и дисциплинарной лексикой имен собственных геообъектов в комплексных исследованиях в контексте определенной дисциплины (в аспекте тектонической, магматической активности, анализа палеоструктур регионов и т.д.) или подхода к рудообразованию (например, флюидно-магматической системы). Среди терминов, обозначающих значительные площади и крупномасштабные объекты, наиболее часто встречаются структурно-тектонические и физико-географические обозначения. В тексте документа территориальная составляющая чаще всего представлена **прямым – именем собственным объекта** или **географическими координатами** (преимущественно для новых, ранее не исследованных, геологических объектов), **относительным** или **косвенным указанием** на геообъект. **Косвенное** указание может быть выражено относительным прилагательным, а также наименованием геологического события или явления. **Относительное указание** чаще обозначается приуроченностью к рекам, озерам, физико-географическим регионам или выражается направлением (например, *к югу от*), расстоянием от инфраструктурных (например, магистрали), техногенных или природных объектов. Прямое, относительное или косвенное указание пространственной привязки тематики документа в тексте может быть снабжено уточнением района, например, указывать на часть физико-географического региона (*центр, центральный, внутренняя часть* или *южный, юг* и т.п.), обозначать положение изучаемой площади относительно бассейнов рек (*верхнее, нижнее* течение), морских акваторий, континентальных водоемов и т.д.

Следует учесть, что в названиях и рефератах текстов по геологии могут присутствовать 1-4 наименования территории, географически соподчиненные в рамках уточнения района исследования, либо объединенные или сопоставленные в едином предмете изучения (например, сравнение месторождений).

Методика и технология автоматизированного извлечения и выделения географической терминологии из текста во многом обусловлены наиболее эффективным разрешением неоднозначностей. На этапе формирования базы данных необходима проверка текстов на ошибки и опечатки. Выделение географической лексики автоматизированным путем происходит в синтезе с разрешением лингвистических не-

однозначностей, выраженных синонимией, гипонимией, омонимией и полисемией терминов, которые необходимо фиксировать в словарях индексирования наряду с другими онтологиями. Более полный перечень онтологических связей дескрипторов в словаре зафиксирован, например, в ГОСТ 7.25-2001. Омонимия, гипонимия, синонимия, ассоциации, указания об обязательном или необязательном использовании термина в составе словосочетаний учитывались и в ранних системах автоматического индексирования [5]. Словосочетания в геообозначениях играют важную роль при разрешении лингвистических неоднозначностей. Корректность картографического отображения пространственной привязки документов также содержит компоненты разрешения неоднозначностей, связанных с территориальным определением документа и дисциплинарным контекстом. К ним можно отнести различного рода неточности, возникающие при «закреплении» на координатной основе пространственной привязки, выраженной сложными соподчиненными отношениями территорий (в том числе с нечеткими границами) или подразумевающей алгоритмы уточнения территории (например, относительной привязки) в режиме ГИС.

Пространственная привязка текстов может зависеть от исторических изменений в наименованиях, а порой и в границах территорий. Эта особенность географической лексики рассматривалась в рамках построения ретроспективного тезауруса для базы данных электронной библиотеки [1, с. 121-128]. Временная зависимость наличия того или иного геообозначения в базе географической лексики и картографического обеспечения характерна и для геологических данных в связи с поэтапным исследованием геологических объектов и разведкой месторождений.

Перечислим основные неоднозначности автоматизированного выделения географической привязки из текстов на русском языке, отмеченные ранее (например, [1, с. 145-151; 254-258]), с дополнениями и особенностями применительно к научным исследованиям рудной геологии.

- Синонимия общепринятых, административных и физико-географических наименований (*Россия – РФ – Российская Федерация, Таймыр – Таймырский полуостров*), геообозначений (*г. – город*), уточнений (*юг – южная часть – Ю.*), а также предметные синонимы (*Биксизакское (Биксизак) рудопроявление*) и аббревиатуры (*ЯАП – Якутская алмазоносная провинция, САХ – Срединно – Атлантический хребет*).

- Омонимия географических названий (совместно между собой, уточнениями и терминами), несущих дисциплинарно-тематическую нагрузку, таких, например, как наименования стратиграфических подразделений (*Золото-кварцевое месторождение Майское (Северная Карелия)*) и *Майское золото-сульфидное месторождение (Чукотка)*); географическое наименование (*месторождение Прогноз*) и описание (*Прогноз ресурсов рудных месторождений*).

- Неточности, связанные с онтологией терминов, входящих в словосочетание (полисемия): *Дара-*

сунское месторождение, Дарасунское рудное поле; медно-никелевые месторождения Печенги, Печенгский рудный узел, Печенгское рудное поле, свинцово-цинковая минерализация Печенгского побережья. Термин, входящий в словосочетание, или его словоформы (синонимы, виды и типы комплексного наименования и т.п.) определяют сам объект и часто его площадь. В связи с этим отметим принятое в геологии различное терминологическое обозначение районирования металлогенических областей для линейновытянутых и иных форм территорий.

- Построение предложений при анализе словосочетаний, несущих пространственную нагрузку. Так, наиболее часто определения по рудно-формационному, генетическому, промышленному типу месторождений или даже по имени собственно месторождения-аналога (*месторождения типа Миссисипи Миссури*) могут «вклиниваться» в пространственное обозначение (*Месторождение золота и сульфида Майское (Чукотка), Ермаковское берtrandит-фенакит-флюоритовое месторождение (Западное Забайкалье, Россия)*). Автоматизированный подход не может ограничиваться только анализом одного, двух соседних слов.

- Предметные ассоциации географических наименований, не относящихся к геологической области. Например, указание местности (населенного пункта), как правило, приурочено к объекту разрабатываемого месторождения.

- Наличие наименований, приуроченных к рекам, озерам, а также к территориям, относящимся к физико-географическому или иному районированию стран, регионов, районов и к другим географическим объектам (*Приэльбрусье, Приладожье, Приазовье, Прикамье, Предчирчингизье, Приколдымье, Приладожье, Средиземноморье (Циркум-Пацифика), Фенноскандия*). Среди такого рода географических обозначений есть устоявшиеся, в том числе имеющие индексы в общепринятых классификациях: *Забайкалье (УДК 571.53/.55), Прибайкалье (571.53/.55), Поволжье (УДК 470.4 и др.), Приморье (УДК 571.63)*.

- Комплексные обозначения территорий (*Африкано-Антарктический регион*).

- Наличие косвенного указания на объекты исследования, выраженное относительными прилагательными, а также наименованиями событий и явлений (*якутские алмазы, т.е. Якутская алмазоносная провинция (ЯАП); Большое Толбачинское трещинное извержение (БТТИ) - вулкан Толбачик*).

- Включение в географическое наименование имен и фамилий известных людей, что осложняет анализ при использовании графематических признаков.

- Наличие географической составляющей в названиях организаций, ведущих исследования, проводящих разработки месторождений или участвующих в цикле обогащения рудного сырья.

- Исторические изменения наименований территорий.

В качестве отличительной особенности при автоматизированном выделении и координатном определении общепринятой географической лексики и лексики предметной можно указать, что термины,

характерные для физико-географического районирования, как правило, рассматриваются в дисциплинарном контексте, обозначая различные, в том числе по площади, геологические структуры. Так, конкретные рудные районы, зоны, узлы обычно включают географическое наименование и название ведущего металла. Перечислим основные неопределенности, связанные с координатным определением в ГИС.

- Относительная географическая привязка рудоносных территорий, приуроченных к рекам, озерам, побережьям и другим физико-географическим объектам, объектам физико-географического районирования, междуречьям (*северо-западное побережье Кольского залива, Западный склон Северного Урала*).

- Тематико-методическая и контекстная зависимости относительной географической привязки. Нередко крупномасштабная пространственная привязка рассматривает группу определенных объектов территории (*золоторудные месторождения Магаданской области, крупные и суперкрупные месторождения*). Такие неоднозначности территориального определения во многом зависят от качества и полноты картографического обеспечения ГИС.

- Полидисциплинарная ландшафтная составляющая, например, территории загрязнений химическими элементами в связи с разработкой месторождения и т.п.

- Присутствие в тексте названий палеоконтинентов и других геообъектов историко-геологической реконструкции.

КАРТОГРАФИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ГИС ДЛЯ ПРОСТРАНСТВЕННОЙ ПРИВЯЗКИ ДОКУМЕНТОВ В ОБЛАСТИ РУДНОЙ ГЕОЛОГИИ

Картографическое отображение пространственной привязки документов и использование ГИС – это приведение к единой координатной основе исследовательской информации по геообъектам различного уровня рассмотрения и пространственной вложенности. Геопривязка текстов может отображаться в виде точечных, линейных или площадных объектов (полигонов) на тематических картах. Глобальный, региональный и локальный уровни рассмотрения, принятые в геологии, подразумевают разномасштабность картографического обеспечения ГИС и возможности перехода от одного способа отображения объекта к другому (от точечного к площадному и т.д.). Например, от рудных (металлогенических) провинций, областей, районов, зон, узлов, месторождений к конкретным рудным телам.

Эффективность, а порой и целесообразность применения анализа пространственной привязки документов с помощью ГИС-сервисов зависят от состава геолого-тематических карт. Картографическое обеспечение ГИС должно содержать базовый состав карт: топографическую основу, рельеф, страноведение, физико-географическое районирование, инфраструктуру. Предметное картографическое обеспечение может включать листы Госгеолкарты, карты структурно-тектонического районирования, неотектоническую

карту разломов, комплекты карт по месторождениям полезных ископаемых (с районированием), карты кайнозойских вулканоплутонических (звенья, провинции, области, районы) и палеовулканических поясов, комплекс карт по подразделениям – возрастным региональным и местным (массивы вулканические, интрузивные, соляные и т.д.), а также карты по техногенным объектам отраслевой и исследовательской инфраструктуры, научно-исследовательским объектам (исследовательские скважины, станции наблюдений и т.п.), карты геологических палео- и современных событий и другие тематические карты.

Использование стандартов – инструмент интеграции картографических данных при распределенном хранении информации на гео-порталах. Международные стандарты в области геоинформатики объединены в общую серию ISO 19100. Отечественные стандарты включают ГОСТ Р 52439-2005 «Каталог объектов местности. Требования к составу» и ГОСТ Р 52573-2006 «Географическая информация. Метаданные».

База данных картографического обеспечения ГИС должна учитывать иерархию объектов без нарушения звеньев классификации, быть тематически логичной и корректно взаимодействовать с системой терминологии автоматического индексирования (например, *Русская плита* и *Русская равнина* - тематически различные термины, территориально совпадающие).

Геологические изыскания могут включать авторский взгляд на геологию региона, что затрудняет пространственное определение текста документа. Это актуально уже на этапе построения картографического обеспечения ГИС при включении в проект тематических карт, как общепринятых, так и генерирующих данные из различных источников. Временная (дата издания карты) и авторская составляющие играют немаловажную роль в корректности картографического обеспечения ГИС для пространственной привязки документов, в особенности для ретроспективных массивов. Авторские карты за разные годы могут быть построены по различным принципам районирования и изучения конкретных рудных объектов. Различия обоснованы подходами, формировавшимися за годы становления предметной области, например, в рассмотрении структурно-тектонического районирования территории, при выделении металлогенических и рудных районов, зон, узлов по структурным признакам, преобладающим рудным формациям и т. п. В связи с авторской составляющей и автоматизированной обработкой текста можно указать на взаимосвязь картографического обеспечения, его пополнения, с собственно-тематической классификацией массива документов исходя из поисковых признаков (терминов), поскольку наибольшую ценность при тематической классификации массива данных представляют уникальные рубрики (словоформы), зачастую авторские. Координатное определение пространственной привязки документов включает компонент вероятностной оценки пространственной принадлежности содержания текста документа (фрагмента текста с гео-

графической лексикой) к территории на карте и рассматривает преимущественно территории с нечеткими границами. Оценка вероятности может учитывать перечисленные выше неопределенности лингвистического характера, таксономию терминологии, авторскую составляющую, а также алгоритм позиционирования территорий, указанных с уточнением, относительно или в неявной форме. Принципом, однозначно определяющим объект, может быть его отражение в утвержденных картах (одобренных Главной редколлегией по геологическому картированию, Научно-редакционным советом Роснедра).

ЗАКЛЮЧЕНИЕ

Среди наиболее значимых составляющих технологии автоматизированного анализа пространственной привязки текстовых сведений с использованием ГИС можно отметить следующие моменты:

эффективность автоматизированного извлечения географических данных из текстов по рудной геологии определяется разрешением лингвистических неоднозначностей, обосновывающих взаимосвязанное ведение тезауруса географической и геолого-географической лексики с тематическими словарями геологических терминов;

корректность отображения географии документов на карте во многом зависит от состава картографического обеспечения ГИС и определяется методикой координатного индексирования данных, которая должна включать подходы к пространственному определению площадных объектов без четких границ, учитывать временную и авторскую составляющие сведений, масштаб, а также предполагать поликритериальную вероятностную оценку принадлежности текста к картографическому обозначению территории.

Комплексный анализ пространственной привязки документов с использованием геоинформационной системы строится на унифицированном картографическом обеспечении и системе терминологии автоматизированной обработки текста, включающей онтологию, правила морфологического анализа и построения словосочетаний, авторскую и временную составляющие.

СПИСОК ЛИТЕРАТУРЫ

1. Труды XIV Всероссийской научной конференции RCDL/2012. Переславль-Залесский, Россия, 15-18 октября 2012 г. – Переславль-Залесский: Изд-во "Университет города Переславля", 2012.
2. Кузьмина Д.А. База данных экологических проектов РАН: пространственные характеристики исследований // Научно-техническая информация. Сер.1. – 2012. – № 2. – С. 12-19.
3. Жижимов О.Л., Мазов Н.А. География и стандарты метаданных для электронных библиотек: содержание, применение, проблемы // РБ. – ИРИО. – 2009. – Том 12, вып. 1. – URL: <http://www.elbib.ru/content/journal/2009/200901/ZM/ZM.ru.html>

4. Куршев Е.П., Сулейманова Е.А., Трофимов И.В. Роль знаний в системах извлечения информации из текста // Электронный научный журнал «Программные системы: теория и приложения» № 3. – Институт программных систем имени А.К. Айламазяна РАН. 2012. – С. 57-70. – URL: <http://www.botik.ru/rented/psta/www/about-magazin/about-mag/about-mag.html>
5. Володин К.И., Гульницкий Л.Л., Пожарский И.Ф. и др. Автоматизированная система научно-технической информации – разработка и эксплуатация. – М.: Финансы и статистика, 2004. – 191 с.

Материал поступил в редакцию 25.10.2013.

Сведения об авторе

КУЗЬМИНА Дарья Александровна – старший научный сотрудник Института геологии рудных месторождений, петрографии, минералогии и геохимии РАН.
e-mail: kuzmida@igem.ru

УВАЖАЕМЫЕ ЧИТАТЕЛИ!

ЦЕНТР НАУЧНО-ИНФОРМАЦИОННОГО ОБСЛУЖИВАНИЯ ВИНИТИ РАН

ПРЕДОСТАВЛЯЕТ КОПИИ ПЕРВОИСТОЧНИКОВ

ВИНИТИ РАН осуществляет обслуживание копиями первоисточников, хранящихся в фонде научно-технической литературы ВИНИТИ, в фондах других библиотек, а также в доступных ВИНИТИ электронных ресурсах.

Фонд научно-технической литературы ВИНИТИ включает более 2 млн изданий по точным, естественным и техническим наукам, в т.ч.:

- отечественные и иностранные периодические и продолжающиеся издания – с 1987 г.;
- отечественные книги – с 1987 г.;
- иностранные книги – с 1991 г.;
- рукописи, депонированные в ВИНИТИ, – с 1962 г.

Заказы на бумажные или электронные копии первоисточников принимает Центр научно-информационного обслуживания (ЦНИО) ВИНИТИ. ЦНИО ВИНИТИ обслуживает коллективных (организации и учреждения) и индивидуальных пользователей.

Формы обслуживания:

- абонементная (на основе договоров и предоплаты);
- разовые заказы (с предоплатой заказа по счету);
- индивидуальная форма обслуживания в читальном зале ЦНИО ВИНИТИ.

На сайте ВИНИТИ (<http://www.viniti.ru>) представлен полный Электронный каталог научно-технической литературы (<http://catalog.viniti.ru>), зарегистрированной в ВИНИТИ с 1994 г. Доступ для просмотра и поиска по Каталогу свободный. Постоянные абоненты ЦНИО ВИНИТИ, имеющие логин и пароль для работы с Каталогом, могут делать заказ копий непосредственно через Каталог.

Услуги по изготовлению копий первоисточников из фондов других библиотек предоставляются только постоянным абонентам. Место хранения первоисточников указывается в Электронном каталоге.

За подробной информацией обращаться по адресу:

125190, Россия, Москва, ул. Усиевича, 20, ВИНИТИ РАН, ЦНИО

Телефоны: 8 (499)155-42-43, 155-42-09, 152-54-59

Факс: 8 (499) 943-00-60

E-mail: cnio@viniti.ru; **URL:** <http://www.viniti.ru>

БАЗА ДАННЫХ ВИНИТИ РАН

ВИНИТИ предлагает к использованию через WWW-сервер (<http://www.viniti.ru>) крупнейшую Федеральную базу отечественных и зарубежных публикаций по естественным, точным и техническим наукам. БД ВИНИТИ РАН генерируется с 1981 г., обновляется ежемесячно, пополнение составляет около 1 млн документов в год. БД ВИНИТИ представлена ретроспективными тематическими фрагментами и единой политематической БД (ретроспектива с 2001 г.), объединяющей все тематические фрагменты БД ВИНИТИ.

БД ВИНИТИ РАН в сети INTERNET

Сервер ВИНИТИ – <http://www.viniti.ru> – обеспечивает on-line доступ к Базе данных ВИНИТИ РАН круглосуточно без выходных.

На основе БД ВИНИТИ РАН предоставляются следующие услуги:

- Диалоговый поиск научно-технической информации в **режиме on-line**;
- **Демо-версия**, позволяющая ознакомиться с основными функциями поисковой системы, составом данных, формами представления документов и получить навыки работы с системой;
- **Поисковые эксперты ВИНИТИ** выполняют тематический поиск по разовым или постоянным запросам, а также окажут **консультационные услуги**.

БД ВИНИТИ РАН на CD-ROM

Любые наборы тематических фрагментов БД ВИНИТИ или их разделов могут быть предоставлены на **CD-ROM в поисковой системе (ИПС) "Сокол"**, обеспечивающей все поисковые функции, доступные в режиме on-line:

- Поиск можно вести в годовом или ретроспективном массиве (за несколько лет сразу) в одном или нескольких тематических фрагментах .
- Поиск по словам и любым словосочетаниям из заглавия, реферата, ключевых слов.
- Использование года, языка, рубрик, шифров тематических разделов БД для уточнения поиска.
- Поиск по словарю, выполняющему функции многоаспектного указателя, в том числе авторского, предметного, источников, индексов МПК, номеров патентных документов и депонированных рукописей и т.д.
- Возможность запоминания запросов для последующего их использования и/или редактирования.
- Чтение документов не только как в РЖ (последовательный просмотр документов одного номера за другим), но и чтение документов нужных тематических фрагментов (разделов) по оглавлению за весь период заказанной ретроспективы.

ИПС "Сокол" является прикладной программой Microsoft Windows.

Любые наборы тематических фрагментов БД ВИНИТИ или их разделов могут быть подготовлены в **коммуникативных форматах ISO-2709, МЕКОФ, txt** на любых видах электронных носителей.

Продукты предоставляются на договорной основе.

Информационная служба БД ВИНИТИ: 125190, Москва, ул. Усиевича 20, ВИНИТИ

Телефон: (499) 155-45-01, 155-45-02, **Факс:** (499) 152-62-31 **e-mail:** csbd@viniti.ru

РОССИЙСКАЯ АКАДЕМИЯ НАУК

Федеральное государственное бюджетное учреждение науки ВСЕРОССИЙСКИЙ ИНСТИТУТ НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ РОССИЙСКОЙ АКАДЕМИИ НАУК

предлагает научным работникам, аспирантам и другим специалистам в области естественных, точных и технических наук, желающим быстро и эффективно опубликовать результаты своей научной и научно-производственной деятельности, использовать способ публикации своих работ через *систему депонирования*.

«Депонирование (передача на хранение) – особый метод публикации научных работ (отдельных статей, обзоров, монографий, сборников научных трудов, материалов научных мероприятий – конференций, симпозиумов, съездов, семинаров) узкоспециального профиля, разрешенных в установленном порядке к открытому опубликованию, которые нецелесообразно издавать полиграфическим способом печати, а также работ широкого профиля, срочная информация о которых необходима для утверждения их приоритета. Депонирование предусматривает прием, учет, регистрацию, хранение научных работ и обязательное размещение информации о них в специальных информационных изданиях».

Подготовка и передача на депонирование научных работ происходит в соответствии с «Инструкцией о порядке депонирования научных работ по естественным, техническим, социальным и гуманитарным наукам» (М., 2003).

Результатом депонирования является публикация информации о депонированных научных работах в информационных изданиях ВИНТИ РАН – Реферативном журнале и аннотированном библиографическом указателе «Депонированные научные работы».

В соответствии с “Положением о порядке присуждения ученых степеней”, утвержденным Постановлением Правительства Российской Федерации от 30.01.2002 № 74 (в ред. Постановлений Правительства РФ от 20.04.2006 № 227, от 02.06.2008 № 424, от 20.06.2011 № 475), научные работы, депонированные в организациях государственной системы научно-технической информации, признаны публикациями, учитываемыми при защите кандидатских и докторских диссертаций.

Подать научную работу на депонирование можно обратившись в Отдел депонирования ВИНТИ РАН по адресу:

125190, Москва, ул. Усиевича, 20.

ВИНТИ РАН, Отдел депонирования научных работ.

Тел.: 8 (499) 155-43-28, Факс: 8 (499) 943-00-60.

e-mail: dep@viniti.ru

С инструкцией о порядке депонирования можно ознакомиться на сайте ВИНТИ РАН:
<http://www.viniti.ru>