

К вопросу о развитии информационной инфраструктуры Сибирского отделения Российской академии наук*

Ю. И. Шокин, А. М. Федотов

Учреждение Российской академии наук

Институт вычислительных технологий СО РАН, Новосибирск, Россия

Новосибирский государственный университет, Россия

e-mail: dir@ict.nsc.ru, fedotov@nsu.ru

Обсуждаются различные аспекты и проблемы развития и поддержки информационной инфраструктуры Сибирского отделения РАН.

Ключевые слова: информационные ресурсы, интеграция ресурсов, информационная инфраструктура, сеть передачи данных, центр обработки данных, информационный центр, доступ к информации.

1. Информационная инфраструктура

Наличие развитой информационной инфраструктуры давно является необходимым условием обеспечения эффективности научных исследований. Интеграция информационных и вычислительных ресурсов в единую среду и организация доступа к ним является одним из важнейших направлений развития современных информационных технологий. Стремительное развитие глобальных компьютерных сетей ведет к изменению фундаментальных парадигм обработки данных вследствие необходимости поддержки и развития распределенных информационно-вычислительных ресурсов. Отметим, что основным направлением работ, определенным правительством РФ, в рамках критических технологий *“является создание инфраструктуры, оборудования, алгоритмического и программного обеспечения для инфокоммуникационных систем и создание взаимоувязанной системы стандартов обработки, хранения, передачи и защиты информации”*.

В настоящее время необходимость разработки механизмов, обеспечивающих функционирование общей информационной инфраструктуры, является приоритетным направлением для задач информационной поддержки научных исследований. Эти вопросы приобретают особую важность для такой организации как Сибирское отделение РАН, в условиях, когда различные группы исследователей, разделенные географически, должны осуществлять совместную работу, обмен данными и знаниями и координировать свои действия с целью оптимизации использования информационно-вычисли-

*Работа выполнена при поддержке РФФИ (гранты № 07-07-00271, 08-07-00229 и 09-07-00277), президентской программы “Ведущие научные школы РФ” (грант № НШ-931.2008.9) и интеграционных проектов СО РАН № 4, 50, 119 и 121.

© ИВТ СО РАН, 2009.

тельных ресурсов, сервисов и приложений. Для крупного территориально распределенного научного центра, каким является Сибирское отделение, — это один из наиболее действенных способов интеграции научных коллективов и применения результатов их исследований в образовании. Острота вопроса наиболее ощущается в крупных интеграционных проектах и при проведении мультидисциплинарных исследований. Здесь информационные технологии играют определяющую роль в разработке базовой инфраструктуры исследований, в том числе и при “публикации” результатов исследований.

Сибирское отделение РАН является региональным объединением научно исследовательских, опытно конструкторских, производственных организаций, а также подразделений, обеспечивающих функционирование инфраструктуры научных центров, расположенных на территории Сибири в семи областях, двух краях и четырех республиках (общая площадь территории около 10 млн. кв. км). Научные центры СО РАН находятся в Новосибирске, Томске, Красноярске, Иркутске, Якутске, Улан-Удэ, Кемерово, Тюмени, Омске, отдельные институты работают в Барнауле, Чите, Кызыле. В состав СО РАН входят более 50 научно исследовательских учреждений, работающих в области физико-математических, технических, химических и биологических наук, наук о Земле, гуманитарных и экономических наук. Примерно половина потенциала Отделения сосредоточена в Новосибирском научном центре.

За годы своего существования информационная инфраструктура СО РАН (сеть передачи данных СПД СО РАН) по числу пользователей и компьютеров, по объемам передаваемых данных, по количеству и качеству накопленных информационных ресурсов, наконец, по разнообразию и качеству предоставляемых услуг СПД СО РАН превратилась в крупнейшую корпоративную научно образовательную сеть России¹.

Характерной особенностью информационной инфраструктуры Отделения является наличие огромного количества аппаратно-технических средств. Последние различаются не только по производителю и характеристикам, но и по платформам и технологиям. Объясняется это несколькими причинами: аппаратура приобреталась в разное время; внедрение производилось разными специалистами, которые использовали разные технологии построения информационной инфраструктуры; инфраструктура развивалась и в географическом плане путем присоединения корпоративных сетей региональных научных центров. Такая историческая многоцветность создает серьезные проблемы для информационной совместимости (итероперабельности) ресурсов и для обеспечения системы безопасности СПД.

Современные требования по организации информационной инфраструктуры ориентированы на создание виртуальной среды и системы необходимых сервисов, требующих создания управляющих информационных центров [1], ответственных за взаимодействие ресурсов, интегрированных в информационную среду. Кроме функций поддержки виртуальной среды, соответствующих сетевых и мультимедийных сервисов, а также управления, синхронизации и диспетчеризации ресурсов СПД, необходимо обеспечивать функции мониторинга за всеми компонентами систем, контроль их параметров, проверку на соответствие, сбор статистики и т. д. Реализация подобных сервисов требует формализации моделей и методов, а также формирования правил (политик) доступа к существующим ресурсам тех или иных проектов.

¹В СПД СО РАН зарегистрировано около 150 организаций абонентов. Только в Новосибирске сеть обслуживает более 40 000 пользователей и насчитывает более 12 000 подключенных компьютеров. Кроме того, в региональных научных центрах Отделения находится еще около 30 000 пользователей.

2. Информационные ресурсы — вечная актуальность

Одним из основных результатов созидательной, социальной и интеллектуальной деятельности человека является создание и накопление информационных ресурсов с целью их дальнейшего использования и недопущения утраты опыта предыдущих поколений. Уровень развития технологий накопления информации и эффективности использования накопленной ранее информации на протяжении всей истории человечества значительно влиял на степень развития производительных сил. Утеря информации приводила к отбрасыванию цивилизации на века назад. Вместе с тем, чтобы эффективно пользоваться накопленной ранее информацией, необходимы специальные инструменты и технологии, при помощи которых могут быть реализованы специальные приемы работы с информацией. Одним из таких приемов является поиск информации.

Наверное, не будет большим преувеличением утверждение о том, что проблема поиска информации является одной из вечных проблем человеческого сообщества. На протяжении многих тысяч лет его представители неустанно находятся в поиске местонахождения *пищи, жилища, пастбища, дороги, сокровища* и т. п. Обобщая задачи поиска, можно сказать, что человечество постоянно находится в поиске *знаний*, в частности, “информации о том, где лежат сокровища”.

Великий аргентинский писатель Хорхе Луис Борхес² в своем эссе “Четыре цикла” писал, что в мировой литературе вечными являются четыре темы:

- 1) падение города;
- 2) возвращение героя;
- 3) поиск;
- 4) самопожертвование бога.

Нетрудно заметить, что наиболее часто встречающейся как в литературе, так и в реальной жизни является третья тема — *поиск*, ибо четвертая тема выходит за рамки обычного человеческого опыта, а две первые проявляются лишь в “минуты мира роковые”.

С появлением новой экономической категории³, какой являются информационные ресурсы, проблема поиска перешла и в эту область. Человеческое сообщество все больше начинает использовать для поиска необходимых знаний информационные ресурсы. Чтобы решить проблему доступа к информации, человечество создало библиотеки как универсальную систему хранения “знаний”, их систематизации и каталогизации [2].

С одной стороны, как отметил британский историк и социолог науки Д. де Солла Прайс⁴, начиная с середины XVIII века любой достаточно большой сегмент науки в нормальных условиях растет экспоненциально, т. е. любые параметры науки, включая объем накопленной информации, за определенный промежуток времени удваиваются (закон экспоненциального роста науки) [3]. С другой стороны, в указанный период времени происходит увеличения числа людей, нуждающихся в научной информации. Речь идет не только о научных работниках (численность которых тоже подчиняется закону экспоненциального роста), но и о представителях многих других профессий умственного труда: инженерах, агрономах, врачах, управленцах и т. д.

²Хорхе Франсиско Исидоро Луис Борхес Асеведо — Jorge Francisco Isidoro Luis Borges Acevedo.

³Информация и информационные ресурсы существовали всегда, но эти ресурсы из-за своей специфичности не рассматривались ранее как отдельная экономическая категория, несмотря на то, что информация всегда использовалась людьми для управления и решения насущных задач.

⁴Дерек Де Солла Прайс — Derek J. de Solla Price.

Любой производственный или научный процесс порождает огромные объемы данных, и по мере того как гигабайты данных превращаются в терабайты работать с ними становится все сложнее. Количество данных когда-нибудь превысит способность компьютеров к их обработке, поэтому необходимо создание новых инструментальных средств и алгоритмов для анализа этих данных. Вместе с тем предъявляются серьезные требования к обеспечению прозрачного доступа и долговременной сохранности “информации”. А в результате вопросы “что хранить?”, “как хранить?” и “как найти?” остаются самыми существенными: без ответа на них все остальные теряют актуальность. Современную технологическую революцию характеризует не центральная роль знаний и технологий, а применение знаний и информации к генерированию знаний и созданию систем, обрабатывающих информацию и осуществляющих передачу “информации”.

Современные информационные технологии предоставляют исследователю мощный аппарат для “манипулирования данными”, а не информацией. Данные, переведенные в электронную форму, приобретают новое качество, обеспечивающее им более широкое распространение и эффективное использование. На первый взгляд, может сложиться впечатление, что развитие информационных технологий уже само по себе способно вывести работу с научной информацией на качественно новый уровень, но, к сожалению, это совсем не так. Современные информационные технологии пока не могут предоставить адекватный аппарат для оперирования с “информацией” и информационными ресурсами [4].

Однако сами по себе данные (как набор битов) без соответствующих описаний или моделей не представляют никакой информационной ценности. Применение информационных технологий должно основываться на использовании различных моделей (феноменологических, информационных, математических и др.). А.А. Ляпунов неоднократно отмечал (см., например, [5]): “нет модели — нет информации”. Для возможности продуктивной работы нужны данные, превращенные в “информацию”, представленную в виде “знаний” — “адекватного отражения действительности в сознании человека в виде представлений, понятий, суждений теорий”.

Как видим, доступ к информации является одной из серьезных проблем, с которой столкнулось современное информационное общество. На заседании Совета по развитию информационного общества, которое проходило под председательством Д.А. Медведева 12 февраля 2009 г., было отмечено, что за последние годы информационные технологии и услуги стали достаточно существенной статьёй российского несырьевого экспорта, достигнув уровня приблизительно в миллиард долларов (с 2000 г. по 2008 г. рост составил 30 % в год). Никакая другая отрасль в РФ не развивалась столь быстрыми темпами. Например, трафик в российском сегменте интернета с 2000 г. возрос в 183 раза. Тем не менее по ключевым показателям в соответствующих международных рейтингах, связанных с информационными ресурсами, Россия занимает 70–80 места. Приведем некоторые факты [6].

— Библиотека Конгресса США за последние 12 лет оцифровала 10 % фондов т.е. 11 млн. материалов (198 млн. долл.). Для оцифровки всего фонда (134 млн. объектов) потребуется не одно десятилетие.

— В Национальной библиотеке Франции в 2006 г. переведено в цифровой формат и размещено на веб-сайтах 80 тыс. текстовых документов. Стоимость оцифровки оценивается в 80 тыс. евро. Фонды французской Национальной библиотеки планируют оцифровывать по 150 тыс. объектов ежегодно.

— Немецкое исследовательское общество для цифровой обработки фондов научных библиотек планирует с 2008 г. предоставлять от 10 до 20 млн. евро.

— В пределах рамочных программ ЕС уже реализованы и реализуются целевые программы, связанные с оцифровкой. Бюджет только одной программы eContentplus (2005–2008), направленной на производство европейского цифрового контента, составляет 149 млн. евро.

Эти проекты направлены на обеспечение пользователей необходимой информацией с целью:

— получения информации через Интернет независимо от времени и места нахождения пользователя;

— существенного повышения оперативности предоставления пользователям необходимой литературы, документов и данных;

— использования машиночитаемых копий для сохранения оригинальных документов и создания страховых массивов документов на случай утраты оригиналов;

— производства работы с электронными документами, которая выходит за рамки простого чтения или просмотра изображения (в том числе редактирование, соединение, добавление, ввод подразделов, перестраивание электронных документов, создание на их основе новых);

— ввода новых форм библиотечного и информационного обслуживания, осуществления доступа значительно большего числа пользователей к документам, имеющимся в библиотеках в ограниченном количестве или в одном экземпляре.

3. Проблемы, связанные с интеграцией ресурсов

Для успешного решения большинства задач, связанных с интеграцией информационных ресурсов, необходимы исследования в следующих направлениях [7]:

— разработка стандартов и моделей (профилей) представления информации и метаинформации;

— построение систем доступа к распределенным и разнородным коллекциям (интероперабельность, масштабируемость, обнаружение релевантной информации, интеграция метаинформации);

— разработка интерфейсов пользователей, визуализация и анализ данных;

— решение вопросов интеллектуальной собственности;

— анализ и обработка естественного языка, изображений, видео- и аудиоданных;

— поддержка многоязыкового доступа к данным и обслуживание данных на нескольких языках;

— разработка мобильных технологий и “интеллектуальных” агентов;

— разработка алгоритмов автоматической классификации информации, методов и средств поиска, каталогизации, индексирования, а также поддержка целостности и непротиворечивости коллекций, безопасность и защита информации.

Несмотря на весьма значительные успехи исследований по многим из перечисленных направлений, сдерживающими факторами формирования единого (виртуального) информационного пространства являются:

— иерархичность информационных систем и ресурсов;

— разнородность ресурсов и программно-технических сред, объединяемых в едином сетевом операционном пространстве;

— распределенность элементов информационной инфраструктуры.

4. Распределенные информационные системы

Как было отмечено, развитие информационных сетей ведет к изменению фундаментальных парадигм работы с информационными ресурсами, в частности, становятся актуальными переход к распределенным ресурсам и создание инфраструктуры для их интеграции в единую информационную систему, обеспечивающую прозрачный доступ к распределенной информации и вычислительным ресурсам.

Любая, в том числе и распределенная, информационная система должна выполнять основные функции, вытекающие из ее основного предназначения, а именно:

- 1) организация хранения информации (организация хранилищ, поддержка систем хранения данных);
- 2) управление информацией (добавление, модернизация, изменение данных);
- 3) управление доступом к информации (контроль исполнения правил регламентации доступа к данным), идентификация данных;
- 4) поиск информации;
- 5) извлечение информации и предоставление ее пользователю в необходимом ему виде;
- 6) визуализация информации в соответствии с требованиями пользователя.

Распределенность и гетерогенность информационных ресурсов налагает следующие дополнительные требования к информационным системам [1]:

- способность систем функционировать в условиях информационной и реализационной неоднородности, распределенности и автономности информационных ресурсов;
- обеспечение интероперабельности, повторного использования неоднородных информационных ресурсов в разнообразных применениях;
- возможность объединения систем в более сложные интегрированные образования, основанные на интероперабельном взаимодействии компонентов;
- осуществление миграции унаследованных систем в новые системы, соответствующие новым требованиям и технологиям при сохранении их интероперабельности;
- обеспечение более длительного жизненного цикла систем.

Информационная инфраструктура любого уровня включает информационные, вычислительные и телекоммуникационные ресурсы. При формировании интегрированной среды основным принципом является функциональная стандартизация или построение функционального стандарта — профиля. Профиль — это согласованный набор стандартов и нормативных документов, в котором в формализованном виде зафиксированы потребности в информационных технологиях, обеспечивающих решение какой-либо задачи или класса задач [8].

С учетом этих требований создание развитой инфраструктуры для представления и обмена метаданными является одним из приоритетных направлений формирования единого информационного пространства и совершенствования современной глобальной информационной сети. В настоящее время многие информационные центры, занимающиеся сбором и распространением метаданных, проявляют активную заинтересованность в организации взаимодействия с целью обмена имеющимися у них ресурсами. Как правило, в основе такой интеграции лежит выработка стандарта на форматы представления метаданных с одновременной унификацией массивов нормативно-справочной информации (разработка профиля информационной инфраструктуры [8]).

При разработке профилей необходимо учитывать постоянное появление новых быстро эволюционирующих типов ресурсов (например, мультимедийные ресурсы, интерак-

тивные сервисы сети, электронные модели объектов, электронные карты, телеконференции, электронные коллекции и т. п.), разработка стандартов для которых в силу их динамической природы и новизны не успевает за темпами развития данных предметных областей. Отметим, что основу интеграции ресурсов составляют технологии работы с метаданными [9], которые:

- обеспечивают механизмы интеграции информационных ресурсов из разных источников сведениями о свойствах этих ресурсов;
- являются источниками сведений о свойствах и содержании информационных ресурсов для механизмов управления данными в информационных системах;
- представляют сведения о системе, ее информационных и других ресурсах для различных приложений и пользователей системы;
- являются источником информации, необходимой для осуществления реинжиниринга информационных систем.

Появление распределенных информационных систем было обусловлено развитием сетей передачи данных, больших корпоративных сетей и глобальной сети Интернет. Задачи распределенных (как и обычных) информационных систем — хранение информации и предоставление ее пользователям в удобном для них виде. Как правило, такие системы могут быть организованы на основе различных технологических решений, направленных на реализацию той или иной парадигмы распределенности. Исходя из основных функций информационных систем, можно рассматривать различные аспекты распределенности:

- 1) распределенное хранение информации (распределенные хранилища, сетевые системы хранения данных, сетевые файловые системы);
- 2) распределенные СУБД и распределенные высокопроизводительные ресурсы;
- 3) управление доступом к распределенным ресурсам и распределенное управление информационными ресурсами;
- 4) поиск информации и информационных ресурсов;
- 5) извлечение информации;
- 6) визуализация информации из распределенных (разнородных) источников в единых пользовательских интерфейсах.

5. Основные функции и технологии

Перейдем теперь к обсуждению технологических решений. Для этого придется вернуться к основным функциям любой информационной системы, сформулированным выше. Несомненно, последняя из указанных функций (визуализация информации в соответствии с требованиями пользователя) проще всего может быть реализована на основе WEB-технологий. Далее, две предпоследние функции (поиск информации в распределенных источниках и извлечение информации из распределенных источников) наиболее просто могут быть реализованы в технологиях, связанных с протоколом Z39.50, так как данный стандарт содержит почти все необходимые для этого компоненты. Во всяком случае, ни один другой стандарт (именно стандарт, причем международный, а не рекомендация) не содержит полного набора необходимых правил, регламентирующих сетевой доступ к базам данных и абстрагирование от конкретных схем и структур данных как для поиска информации, так и для ее извлечения из хранилищ и предоставления пользователю в необходимых ему форматах [10, 11].

Наконец, для реализации функций управления (управление доступом к распределенной информации и распределенное управление информацией) наиболее подходят технологии, основанные на LDAP (упрощенный протокол доступа к каталогам), поскольку именно на его основе сегодня проще всего организовать идентификацию, аутентификацию и авторизацию пользователей в распределенных информационных системах. Немаловажным обстоятельством при этом является тот факт, что LDAP основан на идеологии распределенного хранения информации (деревьев каталогов) на фоне глобальной идентификации всех элементов каталогов, содержит внутри себя определения механизмов и процедур репликаций данных между различными серверами и очень хорошо поддерживается разработчиками прикладного и системного программного обеспечения. Последнее позволяет достаточно просто переходить от локального управления информационными системами и контроля доступа к их ресурсам к распределенному [12].

Следует также заметить, что из вышеперечисленных шести функций первые две, на наш взгляд, эффективнее реализовывать вне парадигм распределенности, обеспечивая тем самым свободу выбора для каждого информационного центра технологической платформы хранения и обработки данных.

Однако здесь необходимо сделать одну существенную оговорку. Приведенная выше схема соответствует некой идеальной технологической модели, каждой из компонентов которой исторически создавался независимо друг от друга для решения вполне определенных собственных задач, зачастую мало связанных с распределенными информационными системами. Каждый из компонентов имеет свои сильные и слабые стороны. Так, академически строгие технологии Z39.50 трудны для реализации в продуктах полупрофессиональных (а таких большинство) разработчиков программного обеспечения и поэтому занимают узкую нишу. Они востребованы лишь там, где без них обойтись невозможно. В большинстве систем универсальность Z39.50 успешно подменяется узкой функциональностью ODBC, JDBC и, совместно с WEB-технологиями, частными решениями, например, на основе PHP. С другой стороны, ограниченность базовых WEB-технологий привела к созданию целой системы технологических надстроек, имеющих конечную цель — построение светлого информационного будущего (“Semantic WEB”). В сторону этого направления сегодня мигрируют многие WEB-ориентированные технологии. Основа миграции — переход на XML и в дальнейшем на сервис ориентированное обслуживание в рамках весьма специфической архитектуры SOAP, основанной на XML-технологиях. Привлекательность такого подхода определяется относительной простотой реализации базовой функциональности, понятной большинству разработчиков информационных систем. Поэтому сегодня мы наблюдаем появление интересных гибридов, сближающих основные технологии построения распределенных информационных систем.

6. Центры обработки данных

Современные тенденции развития информационных технологий в организациях, имеющих разветвленную инфраструктуру и многофункциональные информационные системы, требуют создания экономичных и технологически продвинутых решений, позволяющих наиболее эффективно обрабатывать и хранить информацию, обеспечивать ее доступность и защищенность, а также эффективность использования приложений.

Отметим, что большинство “владельцев” информационных и вычислительных ресурсов, к числу которых относятся и учреждения СО РАН, формируют информационные ресурсы исходя из принципа их приватности и ориентированности на внутреннее использование. В результате большие объемы формально публичной информации труднодоступны или недоступны внешним потребителям, что, в частности, сдерживает мультидисциплинарные исследования и может приводить к курьезным ситуациям. Нередко легче получить научный результат заново, чем узнать о его наличии и получить к нему доступ.

Схожие проблемы существуют и в отношении вычислительных ресурсов, которые вследствие отсутствия надлежащих сервисов являются труднодоступными для большинства потенциальных пользователей и поэтому часто остаются недогруженными, в то время как потенциальным потребителям приходится довольствоваться своими локальными ресурсами, плохо приспособленными для проведения вычислений. Создание единой технологической системы хранения, поиска и обработки информации является одной из важнейших задач интеграции научных исследований. В рамках этой задачи, помимо проблем интеграции первичных информационно-вычислительных ресурсов в единое информационно вычислительное пространство, стоит проблема разработки основных принципов организации таких ресурсов с целью их эффективного использования.

Как уже отмечалось, здесь наиболее экономически адекватным и востребованным решением является создание *Центров обработки данных* (ЦОД), позволяющих аккумулировать мощные вычислительные ресурсы и системы хранения информации, а также резко сократить затраты на обслуживающий персонал и сервисные услуги. Центры обработки данных могут предоставлять информационно-вычислительные ресурсы как непосредственно через развитую систему сервисов, так и в среде “облачных вычислений”. Суть последних состоит в том, чтобы поместить имеющиеся ресурсы в виртуальное “вычислительное облако” так, что доступ к ним можно было бы осуществлять из любого места по мере необходимости. Например, к услугам “облака” могут обратиться удаленный сервер института в пиковый период или рабочая станция в лаборатории, на которой запустили научное приложение, требующее серьезных вычислительных ресурсов. Инфраструктура вычислительного облака обеспечивает гибкое маневрирование ресурсами и их оптимальной загрузки и детального учета объема потребляемых услуг, предоставляемых ЦОД.

В Институте вычислительных технологий создан прототип Центра обработки данных корпоративной распределенной информационной системы, основанной на стандартных протоколах Z39.50, HTTP, LDAP, проведен предварительный этап его опытно промышленной эксплуатации. На базе ЦОД осуществляется доступ к системе хранения данных объемом 70 Тбайт и использования высокопроизводительного кластера для обработки данных. Создано хранилище, которое регулярно пополняется оперативными данными SPOT 2/4 (по прямому каналу из Зап-СибРЦПОД) и включает архивные данные со спутников серии Landsat на территорию РФ за 1982–2002 гг., и создан каталог метаданных, через который осуществляется доступ к информации. При этом отдельные компоненты ЦОД обеспечивают: сервер ZooPARK — доступ к метаданным и документам в соответствии с протоколом Z39.50 и HTTP, Sun Java System Directory Server — доступ к каталогам по протоколу LDAP, Sun Java System Application Server — выполнение JSP- и других приложений в соответствии с J2EE, DSpace — управление

репозитарием цифровых объектов (полнотекстовые документы, мультимедийные объекты, изображения и т. п.), PostgreSQL — исполнение функций реляционной СУБД.

Создание прототипа ЦОД еще раз показало необходимость разработки профиля информационной инфраструктуры СО РАН для поддержки фундаментальных исследований. Отметим, что при построении профиля информационной инфраструктуры следует использовать комбинацию множества базовых стандартов и/или отдельных профилей. Все множество ресурсов, составляющих инфраструктуру, разделяется на три категории:

- 1) телекоммуникационные ресурсы;
- 2) ресурсы коллективного доступа (центры хранения данных, электронные библиотеки, вычислительные ресурсы и т. п.), доступные любому исследователю;
- 3) частные ресурсы организаций, в которых проводятся фундаментальные исследования, эти ресурсы используются, как правило, в конкретных проектах либо группой лиц, занятых конкретным исследованием.

Профиль информационной инфраструктуры следует строить как профиль общего назначения, в который как подмножества входят отдельные части профилей конкретного применения — профиль организации.

В основу построения профиля следует положить методiku, основанную на принципе декомпозиции и включающую несколько стадий последовательного перехода от анализа предметной области до создания собственно профиля. Декомпозиция может выполняться применительно к различным моделям информационной среды, что позволяет установить соответствие между элементами модели и элементами профиля. При построении профиля используется модель открытых систем с выделением в ней групп служб профиля. На завершающем этапе выявляются службы, которые стандартизируются в соответствии с рекомендациями ISO/IEC.

В заключение остановимся на двух моментах.

1. Во-первых, отметим, что создание информационной инфраструктуры и связанных с ней систем управления ресурсами и информационной безопасности требует значительных людских и материальных затрат.

2. Во-вторых, цель создания информационной инфраструктуры — обеспечение конечного пользователя необходимой информацией и информационно-вычислительными ресурсами. Пользователя, как правило, не интересует, как устроена технологическая “кухня” системы. Более того, чем меньше эта “кухня” заметна пользователю, тем лучше построена та или иная система. Следует думать, что использование проверенных временем и практикой типовых решений и международных стандартов позволит обеспечить эту технологическую “прозрачность”.

Список литературы

- [1] Шокин Ю.И., Федотов А.М., Жижимов О.Л. Технологии распределенных информационных систем // Современные информационные технологии для научных исследований: Материалы Всерос. конф. Магадан, 2008. С. 18–21.
- [2] Федотов А.М., Баракнин В.Б. Проблемы поиска информации: История и технологии // Вестник НГУ. Серия: Информационные технологии. 2009. Т. 7, вып. 2. С. 3–17.

- [3] ПРАЙС Д. Малая наука, Большая наука: Пер. с англ. // Наука о науке. М.: Прогресс, 1966. С. 281–385.
- [4] ФЕДОТОВ А.М. Парадоксы информационных технологий // Вестник НГУ. Серия: Информационные технологии. 2008. Т. 6, вып. 2. С. 3–14.
- [5] ЛЯПУНОВ А.А. О соотношении понятий материя, энергия и информация // Проблемы теоретической и прикладной кибернетики. Новосибирск: Наука, 1980. С. 320–323.
- [6] СЮНТЮРЕНКО О.В., БОРИСОВА Л.Ф. Электронные библиотеки и проблемы интеграции информационных ресурсов в гетерогенном сетевом пространстве // Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса: материалы конф. Электрон. дан. М.: ГПНТБ России, 2009 (CD-ROM). ISBN 978-5-85638-132-9. № гос. регистрации 0320900806.
- [7] ЖИЖИМОВ О.Л., ФЕДОТОВ А.М., ЧУБАРОВ Л.Б., ШОКИН Ю.И. Технология создания распределенных информационно-вычислительных ресурсов СО РАН // Тр. I Междунар. конф. САИТ-2005. “Системный анализ и информационные технологии”. Переславль-Залесский, 2005. Т. 2. 161–165.
- [8] IEEE Std 1003.23-1998, IEEE Guide for Developing User Organization Open System Environment (OSE) Profiles.
- [9] TASK Force on Metadata. Summary Report // American Library Association. 1999. Vol. June.
- [10] ШОКИН Ю.И., ФЕДОТОВ А.М., ГУСЬКОВ А.Е. и др. Электронные библиотеки — путь интеграции информационных ресурсов Сибирского отделения РАН // Вестник КазНУ. Спецвыпуск. 2005. № 2. С. 115–127.
- [11] ЖИЖИМОВ О.Л., ТУРПАНОВ А.А., ФЕДОТОВ А.М. Корпоративный каталог СО РАН // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Тр. Восьмой Всерос. науч. конф. (RCDL’2006). Ярославль, 2006. С. 226–230.
- [12] БАРАХНИН В.Б., ЖИЖИМОВ О.Л., СТЕПАНОВ Ю.Ю., ФЕДОТОВ А.М. LDAP-каталог организации как ядро корпоративной распределенной информационной системы // Инновационные недра Кузбасса. IT-технологии: Сб. науч. тр. Кемерово: ИНТ, 2008. С. 226–232.

Поступила в редакцию 7 января 2009 г.