

ЖИХАРЕВ Александр Павлович – кандидат технических наук, заместитель генерального директора ГУП «Московский социальный регистр», руководитель Службы ведения общегородских классификаторов и Единого реестра информационных ресурсов и систем города Москвы

ВОПРОСЫ ПРИМЕНЕНИЯ СИСТЕМ КЛАССИФИКАЦИИ И СПРАВОЧНОЙ ИНФОРМАЦИИ ПРИ УЧЕТЕ И АНАЛИЗЕ ИНФОРМАЦИОННЫХ РЕСУРСОВ

Введение

Определяющими предпосылками эффективного управления информационными ресурсами являются: организация учета существующих ресурсов, закрепление их в собственности (материальной и/или интеллектуальной), в оперативном управлении и хозяйственном ведении, определение правомочий владельцев [1]. Полноценный учет любых объектов невозможен без использования элементов систематизации этих объектов. Основные составляющие систематизации – унификация (нормализация), классификация и кодирование. Существует множество разновидностей информационных ресурсов и, как следствие, множество подходов к их классификации. В настоящей работе рассмотрен ряд систем учета государственных и муниципальных информационных ресурсов на основе реестров информационных ресурсов и систем различного уровня (федеральный, уровень субъектов РФ). Основное внимание уделено использованию в системах учета классификаторов, справочников, словарей различных типов.

Существующие системы учета и регистрации информационных ресурсов и подходы к их классификации

Государственный регистр баз и банков данных

На Федеральном уровне учет информационных ресурсов осуществляется в соответствии с постановлением Правительства РФ “О государственном учете и регистрации баз и банков данных” от 28.02.1996 г. № 226. Учет производится в Государственном регистре баз и банков данных (далее по тексту – Государственный регистр). Головной организацией по ведению Государственного регистра является ФГУП «Информрегистр», который формирует, в частности, электронный каталог Государственного регистра, издает каталог “Базы данных России“ (www.inforeg.ru). Структура каталога включает разделы: “Систематический указатель баз данных“ и “Указатель владельцев баз данных“. В первом разделе используется группировка сведений по тематическому признаку в соответствии с Государственным рубрикатом научно-технической информации (ГРНТИ) (далее по тексту – рубрикат). Систематический указатель включает 70 группировок рубрикатора. Использование такого количества группировок вполне оправдано, так как в Государственном регистре содержится значительное количество самых разных информационных ресурсов. Эффективность применения рубрикатора в других реестрах, например, на уровне субъектов РФ, городов и т.п. не является очевидной ввиду меньших объемов и тематического разнообразия объектов, а также по причине сложности процесса индексации.

В Государственном регистре применяются также локальные справочники (перечни).

1. *Вид документа.* Справочник содержит 37 позиций: архивная информация; нормативные и законодательные документы; управленческая информация; статистическая информация; справочная информация и т.д.

2. *Назначение базы данных.* Справочник включает позиции: архивная; исследовательская; коммерческая; массовая; НТИ; официальная; проектно-конструкторская; производственно-техническая; справочная; управленческая; учетно-статистическая; учебная; финансовая.

3. *Тип базы данных.* Справочник содержит следующие типы: библиографическая; полнотекстовая; фактографическая; гипертекстовая; библиографическая с рефератом; документально-фактографическая; лексикографическая; АРМ; ГИС; мультимедиа; обучающая система; картографическая; комбинированная.

Фиксация данных параметров в описаниях информационных ресурсов и применение классификации баз данных в соответствии со справочниками 1,2,3 вызывает сомнения. Справочники пересекаются между собой. Трудности отнесения объекта учета к конкретному виду очевидны. Опыт показывает, что данные позиции при формировании исходных регистрационных документов вызывают у заявителей существенные трудности.

4. *Вид конфиденциальной информации.* Справочник включает позиции: коммерческая тайна, следственная тайна, профессиональная тайна, служебная тайна, персональные данные, патентная заявка, без ограничений. Такое разделение имеет определенные недостатки. Во-первых, законодательные, нормативные документы не содержат однозначных определений всех этих понятий. Во-вторых, как показывает опыт, при регистрации информационных ресурсов конкретный вид ограничений по доступу к информации не подтверждается какими – либо документами, что снижает достоверность данных. Поэтому более эффективным представляется разделение ресурсов на ресурсы открытого (свободного) доступа и ресурсы ограниченного доступа. Дополнительная расшифровка «ограниченного доступа» может носить справочный характер. Формирование статистики по данному параметру весьма важно, так как задача повышения открытости информационных ресурсов является одной из актуальных.

5. *Вид носителя информации при ее поставке (распространении).* Использование такого рода классификации информационных ресурсов вряд ли оправдано, т.к. в настоящее время практически любой владелец имеет возможность поставки ресурса с использованием различных типов носителей. Однако в качестве характеристики информационного ресурса данный параметр необходим.

6. *Период обновления.* Данный признак полезен при описании регистрируемого ресурса, так как дает представление об уровне и трудоемкости актуализации. Классификация ресурсов по данному признаку с целью формирования соответствующих группировок не имеет конструктивной цели.

7. *Форма собственности владельца.* В перечне приводятся позиции: государственное учреждение; иностранная фирма; предприятие, организация с привлечением иностранного капитала; коммерческая структура; муниципальная форма собственности; общественная организация; региональная форма собственности; предприятие со смешанным капиталом; федеральная форма собственности; владелец – физическое лицо; частное предприятие. Представляется, что в приведенном перечне смешаны такие нормы, как «форма собственности» и «организационно-правовая форма». Для первой нормы существует Общероссийский классификатор ОКФС, для второй – Общероссийский классификатор ОКОПФ. Строго говоря, для фиксации форм собственности владельцев, равно как и других юридических лиц – субъектов реестров, рекомендуется использовать ОКФС.

8. *Тип используемой СУБД.* Перечень содержит свыше 100 наименований современных универсальных системных средств хранения и обработки данных. Наличие данного параметра при описании объектов учета можно считать обоснованным. Статистические данные, которые можно получить на основе классификации (группировки) ресурсов по данному признаку, позволят определить уровень типизации используемых системных средств, проконтролировать права использования системных средств (наличие лицензий и т.п.).

9. *Тип предоставляемых услуг (избирательное распространение информации, ретроспективный поиск, тиражирование БД, поставка фрагмента БД).* Аналогичный состав информационных услуг, предоставляемых на основе информационного ресурса, должен содержаться в системе учета. Приведенный перечень может быть расширен. Классификация объектов учета по данному признаку вряд ли возможна и целесообразна, так как каждый информационный ресурс, как правило, характеризуется несколькими типами информационных услуг.

10. *Источник финансирования создания (ведения).* Данный признак при учете и классификации информационных ресурсов следует признать необходимым для получения объективных данных использования (соотношения) бюджетов и иных источников для создания и ведения информационных ресурсов.

11. *Тип организации владельца.* Не приводя здесь полный перечень позиций справочника, отметим, что он сформирован по близкому к отраслевому принципу (...торговые и сбытовые организации, транспортные предприятия, учебные заведения, органы государственной власти и т.д.). Нет достаточных оснований для эффективного использования данного справочника, например, для региональных государственных реестров, т.к. в них основной тип организации владельца – органы государственной власти региона.

12. *Справочники: тип ЭВМ, тип ОС, наличие удаленного доступа к БД.* Данные справочники необходимы для описания информационного ресурса любого реестра. С их использованием должна формироваться соответствующая статистика.

13. *Язык БД.* В рамках ведения Государственного регистра данный признак необходим. Для региональных государственных реестров – реестров, содержащих в основном информационные ресурсы органов власти, - не является актуальным.

Реестр государственных информационных ресурсов Санкт-Петербурга

Порядок формирования и ведения Реестра государственных информационных ресурсов Санкт-Петербурга регламентируется Распоряжением Администрации Санкт-Петербурга от 17 октября 2001 года № 1024 –ра «Об утверждении Временного положения об информационном взаимодействии исполнительных органов государственной власти Санкт-Петербурга между собой и с органами местного самоуправления в Санкт-Петербурге» (www.iac.spb.ru). Документы не содержат прямых упоминаний об использовании каких-либо систем классификации информационных ресурсов. Однако из описания состава учитываемых ресурсов следует, что они классифицируются как:

а) государственные информационные ресурсы Санкт-Петербурга;

б) информационные ресурсы органов местного самоуправления.

В свою очередь, государственные информационные ресурсы подразделяются на:

а.1) базовые государственные информационные ресурсы Санкт-Петербурга;

а.2) производные государственные информационные ресурсы Санкт-Петербурга.

Базовые ресурсы содержат документированную информацию обо всех субъектах и объектах, находящихся на территории Санкт-Петербурга (типовые по составу и структуре кадастры, реестры, классификаторы и др.). Производные ресурсы формируются путем добавления функциональных признаков субъектов и объектов к их описанию в существующих базовых ресурсах или посредством обработки информации, описывающей объекты и субъекты, находящиеся на территории Санкт-Петербурга. Перечень и состав базовых и производных информационных ресурсов определяются потребностями владельцев и пользователей, а также могут регламентироваться соответствующими нормативными документами Российской Федерации и Санкт-Петербурга. Из структуры паспорта информационного ресурса следует, что регистрируемые объекты (ИР) подразделяются также на:

единые централизованные ИР, содержащие все показатели;

централизованные ИР, содержащие обобщенные показатели;
 централизованные ИР, содержащие подмножества показателей распределенных ИР;
 распределенные ИР.

По степени конфиденциальности информации информационные ресурсы разделяются на общедоступную информацию; конфиденциальную информацию I категории; конфиденциальную информацию II категории. Присвоение категорий осуществляется в соответствии с Положением о категорировании конфиденциальной информации органов Администрации Санкт-Петербурга, утвержденным вице-губернатором-руководителем Канцелярии губернатора Санкт-Петербурга от 20.03.2001 г.

Тот факт, что степень конфиденциальности данных, содержащихся в информационных ресурсах, устанавливается администрацией субъекта, безусловно, является положительным. Однако корректность выделения категорий нуждается в дополнительном анализе. Паспорт информационного ресурса содержит также следующие реквизиты, по которым возможно осуществление группировок регистрируемых объектов:

тип используемых при формировании (ведении) ИР классификаторов (справочников) (общероссийский, общегородской, ведомственный, внутрисистемный);

наличие обмена с внешними информационными системами (да/нет);

способы обмена (Internet, Intranet, магнитные носители);

периодичность обмена (по запросу, ежедневно, раз в квартал и т.п.);

наличие потребности в информационных ресурсах других организаций (да/нет).

Реестр государственных и муниципальных информационных ресурсов и систем Московской области

Правовой основой функционирования Реестра государственных и муниципальных информационных ресурсов и систем Московской области являются следующие документы: Постановление Губернатора Московской области от 30.06.99 № 224-ПГ «О Реестре государственных и муниципальных информационных ресурсов и систем Московской области»; Закон Московской области от 11.01.2001 № 21/2001 – ПГ «О создании органами государственной власти Московской области информационных ресурсов и информационных систем и их использовании»; Постановление Губернатора Московской области от 19.05.2001 № 145-ПГ «О порядке государственного учета информационных ресурсов и информационных систем Московской области».

В составе целей создания Реестра отмечено решение комплекса аналитических задач по выявлению основных тенденций в развитии государственных и муниципальных информационных ресурсов и систем Московской области. Достижение данной цели возможно при наличии объективных статистических данных, полученных по определенным классификационным группировкам информационных ресурсов и систем в реестре. Из текста Положения о данном реестре следует, что в качестве таких классификационных группировок предусматриваются: вновь создаваемые информационные ресурсы и системы; действующие информационные ресурсы и системы; ликвидированные информационные ресурсы и системы.

Кроме этого, включаемые в состав реестра информационные ресурсы должны в соответствии с Положением разделяться по праву собственности на:

информационные ресурсы РФ;

информационные ресурсы Московской области;

информационные ресурсы муниципальных образований на территории Московской области;

информационные ресурсы, находящиеся в общей собственности указанных выше субъектов,

а также по своему назначению на: реестровые ресурсы и объектно-ориентированные ресурсы.

В заявлении на регистрацию предусмотрены коды заявителей (владельцев) по: ОКПО, ОКОПФ, ОКОНХ, ОКДП, ОКФС. При формировании данных об объекте регистрации используется локальный классификатор источников финансирования работ по созданию и актуализации информационного ресурса.

Формы собственности информационного ресурса и/или системы в рассматриваемом реестре кодируются по классификатору ОКФС, что вызывает сомнения. Например, в ОКФС отсутствуют такие позиции, как: собственность юридических лиц; смешенная (совместная, общая) собственность в различных сочетаниях (РФ, субъектов РФ, муниципальных образований). Таким образом, по ОКФС не может быть закодирована, например, общая собственность Московской области и муниципального образования, расположенного на территории Московской области. Видимо, отчасти и поэтому, в статистике по реестру отсутствуют данные о количестве зарегистрированных ИС и ИР, находящихся в «общей» собственности. Вместе с тем, фиксация в реестрах конкретных собственников (обладателей исключительных прав) на БД [2] является необходимой. Регистрируемые информационные ресурсы классифицируются также по типам в соответствии с локальным классификатором типов информационных ресурсов [3]. Классификатор содержит 51 позицию. Для сравнения – Систематический указатель баз данных каталога «Базы данных России» содержит около 70 позиций. Учитывая объем реестра (338 объектов), классификатор, даже с учетом естественного увеличения объема реестра, кажется слишком подробным, что может снизить эффективность статистики.

Одна из поставленных целей данного реестра - унификация и стандартизация программных и технических решений в сфере информатизации. Данная цель достижима при наличии статистической информации об используемых решениях и средствах. Для этого могут использоваться следующие реквизиты из состава предусмотренных в реестре данных об объекте регистрации: вид носителя информации; ограничения по доступу к информации; тип ЭВМ; тип используемых системных средств; формат данных; протоколы доступа; тип актуализации данных; период актуализации данных; наличие сертификатов. Формирование требуемых

группировок наиболее эффективным образом может быть достигнуто с использованием соответствующих справочников.

Единый реестр информационных ресурсов и систем города Москвы

Нормативной базой создания Реестра являются следующие основные документы: Закон г.Москвы от 24 октября 2001 г. № 52 “Об информационных ресурсах и информатизации города Москвы”; постановление Правительства Москвы от 24 июня 2003 г. № 496-ПП «О Единой системе классификации и кодирования и Едином реестре информационных ресурсов и систем города Москвы». Среди задач и функций Службы ведения реестра указаны, в частности, следующие [4]: формирование и представление статистических и аналитических данных о состоянии информационных ресурсов и систем г. Москвы оператору реестра для осуществления процедур управления процессами информатизации в городе; анализ и прогнозирование количественных и качественных характеристик информационных ресурсов и систем; информационно-поисковые услуги. Очевидно, что эффективное выполнение этих задач должно опираться на применение элементов унификации и классификации информации реестра.

Специальных требований к системам унификации и классификации информационных ресурсов и систем данного реестра анализируемые документы не содержат. Прямо указано лишь на разделение объектов учета на регистрируемые временно (предварительно) и постоянно. Предварительная (временная) регистрация осуществляется после принятия решения о создании системы и (или) формировании информационного ресурса, постоянная – после ввода в действие (в эксплуатацию). Формирование группировок и статистики по данному признаку является необходимым для проведения мониторинга состояния реестра и выработки соответствующих воздействий. В заявлении на регистрацию в реестре предусмотрены коды общероссийских классификаторов заявителей: ОКПО, ОКВЭД, ОКФС, ОКОПФ.

По состоянию на 12.05.2005 года в соответствии с классификатором ОКФС 82% заявителей реестра представляют собственность города Москвы - субъекта РФ. В соответствии с ОКОПФ 63% составляют заявители – учреждения; 29% - унитарные предприятия, основанные на праве хозяйственного ведения. Права собственности на информационные ресурсы в данном реестре фиксируются прямым указанием собственника в паспорте на информационный ресурс.

Реестр содержит более 20 локальных словарей (справочников, перечней) (www.mosreestr.ru):

1. Источники финансирования; 2. Качественный эффект; 3. Тип субъектов реестра; 4. Назначение ИР (ИС); 5. Наименования компонентов ИС; 6. Тип рабочих станций и терминалов; 7. Тип серверов; 8. Тип операционных систем; 9. Периодичность обновления БД; 10. Режим предоставления данных; 11. Средства защиты данных; 12. Стадии ИР (ИС); 13. Статус классификатора (справочника, словаря); 14. Категория доступа к информации; 15. Тип универсальной СУБД; 16. Тип базы данных; 17. Тип нормативно-справочных данных; 18. Тип нормативных, распорядительных документов; 19. Тип документа-основания передачи/получения/использования данных; 20. Тип носителя данных; 21. Языки программирования.

Словари с номерами: 2, 4, 6, 9, 11, 13, 14, 16, 21 используются для нормализации (унификации) информации в реестре. Словари с номерами: 1, 7, 8, 11, 12, 14, 15, 21 используются для нормализации информации, организации поиска, формирования группировок, выборок, статистики. Остальные словари носят справочный характер. Словари применяются в процессе формировании регистрационных форм заявителями, также при интерактивном вводе информации в базу данных и ее контроле персоналом Службы ведения реестра.

Выше говорилось о сомнениях в эффективности прямого использования рубрикатора ГРНТИ для классификации информационных ресурсов в региональных реестрах, вызванных, в частности, трудностями индексации информации. Однако эксперимент, проведенный на базе Реестра информационных ресурсов и систем города Москвы, показал возможность обойти этот трудоемкий процесс. Суть заключается в следующем. Информационные ресурсы Москвы помимо учета в реестре субъекта Российской Федерации регистрируются в Государственном регистре баз и банков данных. По соглашению с НТЦ «Информрегистр» Москва получает от НТЦ актуализированный массив данных, содержащих информацию обо всех зарегистрированных в Государственном регистре ресурсах организациями, расположенными на территории города Москвы, включая и ресурсы, являющиеся собственностью города Москвы. Данный массив уже проиндексирован в соответствии с ГРНТИ. Искомый проиндексированный массив получается путем выделения в автоматическом режиме из общего набора данных перечня ресурсов, являющиеся собственностью города Москвы. Сравнение полученного систематизированного массива с каталогом «Базы данных России» показал, что информационные ресурсы города Москвы распределились по 28 рубрикам, что составляет 40% от общего числа рубрик, используемых в Российском каталоге. Данный результат представляется достаточно оптимистичным с точки зрения использования данного типового подхода для классификации информационных ресурсов региональных реестров, для выпусков каталогов, справочников, указателей, формирования тематических подборок.

Выводы, рекомендации

Рассмотренные системы учета и регистрации информационных ресурсов и систем на базе функционирующих реестров информационных ресурсов имеют много общего как по содержанию организационно-методического обеспечения, так и по составу учитываемых данных. Все проанализированные реестры в той или иной степени используют методы систематизации и классификации информации. Однако можно констатировать отсутствие единого подхода к классификации информационных ресурсов на федеральном и

региональном уровнях. Единые требования к региональным реестрам информационных ресурсов могут быть разработаны путем интеграции нормативных документов, рассмотренных в работе реестров. Можно констатировать недостаточное использование информации реестров для формирования статистических и аналитических данных, в частности, с применением классификаторов и справочников.

В процессе учета и регистрации ИР должны использоваться системы классификации различных уровней и типов: Общероссийские классификаторы; тематические рубрикаторы информации (в частности ГРНТИ); универсальные классификаторы типов ИР; локальные справочники (словари, перечни).

Из Общероссийских классификаторов рекомендуется применять:

- ОКПО - для контроля и отслеживания действующих наименований организаций – заявителей (владельцев), других субъектов реестров, которые подвержены изменениям;

- ОКОПФ, ОКФС - для учета характеристик юридических лиц – субъектов реестров и формирования соответствующей статистики.

ГРНТИ целесообразно использовать (по предложенной в работе технологии) при выпуске справочников и каталогов, для формирования различных тематических подборок. На базе ГРНТИ может быть создан укрупненный рубрикатор с ограниченным количеством позиций, могут разрабатываться другие рубрикаторы, отражающие специфику конкретного реестра (региона).

Выбор состава локальных справочников (словарей) зависит от целей и задач реестров и состава реквизитов паспортов регистрируемых информационных ресурсов.

Универсальные классификаторы типов ИР применимы для формирования группировок и статистической информации об ИР.

В [5] приведен текст классификатора типов информационных ресурсов, который может быть взят за основу для классификации информационных ресурсов субъектов РФ.

Для классификации информационных ресурсов в региональных реестрах необходим иерархический классификатор органов власти, управления, подведомственных предприятий и организаций региона. Важность получения статистики распределения информационных ресурсов по органам власти и управления продиктована тем обстоятельством, что именно эти органы в основном являются субъектами, осуществляющими правомочия собственников государственных информационных ресурсов.

В ряде реестров помимо реализации чисто регистрационных функций выполняются (планируются) другие функции: учет и мониторинг прав на ИР, учет затрат на создание и ведение ИР, учет информационных услуг, оказываемых на основе ИР, и ряд других. Для таких реестров требуются дополнительные справочники и перечни: информационных услуг; типов затрат; типов прав и т.д. Для осуществления мониторинга процесса регистрации в реестрах целесообразно также использовать справочник этапов и стадий регистрации.

Важным является также придание используемым в реестрах классификаторам и справочникам определенного статуса путем утверждения их на уровне регионов и включения в Единые системы классификации и кодирования технико-экономической и социальной информации регионов [4].

Литература

1. Информационные ресурсы России. Национальный доклад. Государственный комитет Российской Федерации по телекоммуникациям. М., 1999.

2. «О правовой охране программ для электронных вычислительных машин и баз данных». Закон Российской Федерации от 23 сентября 1992 г. № 3523-1.

3. Методические указания о порядке внесения заявителями документов для регистрации информационных ресурсов и систем в Реестре государственных и муниципальных информационных ресурсов и систем Московской области (www.mintransmo.ru).

4. «О Единой системе классификации и кодирования и Едином реестре информационных ресурсов и систем города Москвы». Постановление Правительства Москвы от 24 июня 2003 г. № 496-ПП.

5. Жихарев А.П. Использование систем классификации при учете информационных ресурсов// Компьютер. – 2005. - июль/август.