KISELEV M.YU.

DATABASE “ELECTRONIC INVETNTORY RAS ARCHIVE”: STAGES OF CREATION AND DEVELOPMENT PERSPECTIVES

М.Ю. КИСЕЛЕВ

БАЗА ДАННЫХ «ЭЛЕКТРОННЫЕ ОПИСИ АРХИВА РАН»: ЭТАПЫ СОЗДАНИЯ И ПЕРСПЕКТИВЫ РАЗВИТИЯ

Annotation / Аннотация

In terms of automatic scientific and informative system creation, there is an “Accountancy and catalogue” database developed, as a part of information system the Archive of RAS. To implement the possibility of studying archive inventories using the information system the Archive of RAS, there was set a goal of creating “Electronic inventory of the Archive of RAS”. The stages of creation and development perspectives of “Electronic inventory RAS Archive” are considered.

В статье рассматривается проблема создания автоматизированного научно-справочного аппарата в Архиве РАН, разработка базы данных «Учет и каталог» в составе информационной системы «Архив РАН». Для реализации возможности изучения пользователями архивных описей архива через информационную систему «Архив РАН» была поставлена задача создания базы данных «Электронные описи Архива РАН». Рассмотрены этапы создания и перспективы развития базы данных «Электронные описи Архива РАН»

Keywords / Ключевые слова

Database, electronic inventory, the Archive of RAS, information technologies, automatic scientific and informative system. База данных, электронные описи, Архив РАН, информационные технологии, автоматизированный научно-справочный аппарат.

КИСЕЛЕВ Михаил Юрьевич – руководитель Центра учета и обеспечения сохранности документов Архива Российской академии наук, кандидат исторических наук, г. Москва; 8-499-129-37-66, 8-495-613-68-97, kiss_RAN@mail.ru

Архив Российской академии наук, созданный в 1728 г., остается крупнейшим и старейшим хранилищем архивных документов по истории Российской академии наук, ее учреждений и организаций, отечественной и зарубежной науки и культуры, российских и иностранных ученых. Происходящие в России процессы, связанные с развитием информационного общества, реформированием Российской академии наук и ее учреждений, повышением интереса к истории науки, ставят задачу организации

высокоэффективной информационной системы поиска документной информации. Одним из фундаментальных направлений своей деятельности Архив РАН ставит обеспечение ретроспективной информацией, имеющей важную научную и социокультурную значимость, учреждений Российской академии наук и исследователей.

В решении этой задачи наиболее эффективным для Архива РАН являлось создание автоматизированного научно-справочного аппарата (НСА), который, реализуя все функции традиционной системы научно-справочного аппарата, предоставляет по сравнению с ней более широкие возможности оперативного и многоаспектного поиска архивных документов и представления его результатов. В Архиве РАН сложилась система научно-справочного аппарата: путеводитель, справочники, описи, каталоги и картотеки, базы данных к отдельным документальным комплексам. Однако она не позволяла решать ряд проблем: низкая оперативность поиска и предоставления архивной информации; отсутствие возможности одновременной работы нескольких пользователей с одним и тем же документом; использование трудозатратных традиционных («ручных») способов описания документов и ведения картотек и каталогов; отсутствие оперативной возможности получения информации о текущем состоянии учета, научного описания и сохранности документов архива.

Начало использованию информационных технологий для создания автоматизированного научно-справочного аппарата в Архиве РАН было положено в 2004-2006 гг. при разработке многофункциональной базы данных «Учет и каталог», которая была разработана при финансовой поддержке Российского гуманитарного научного фонда (РГНФ). В основу программного обеспечения был положен Общий Международный стандарт описания архивных документов ISAD(G), утвержденный Специальной комиссией по стандартам Международного совета архивов в 1993 г. Международный стандарт закрепил принцип многоуровневой структуры данных (фонд - опись - дело - документ) в научном описании архивных документов, установил органическую связь между архивной деятельностью и основами теории информатики на всех этапах: от создания информации до ее использования.

База данных представляет собой структурированную совокупность архивных данных, сосредоточенных в двух разделах: «Учет» и «Каталог». Раздел «Учет» состоит из четырех описательных уровней: фонды - описи - дела -документы. На каждом описательном уровне заложены карты (формы) с унифицированными форматами метаданных - информационными полями, которые являются описательными элементами данного уровня. В распределенной структуре базы данных на описательном уровне «Описи» заложена функция создания и осуществляется хранение графических файлов с оцифрованными изображениями традиционной описи фонда на бумажной основе. На описательных уровнях «Дела» и «Документы» заложена функция создания и осуществляется хранение электронного фонда пользования - полнотекстовых (факсимильных) изображений документов. Карты (формы) с унифицированными форматами метаданных двух нижних уровней информационной структуры «Дела» и «Документы» интегрированы в раздел «Каталог».

Раздел «Каталог» предназначен для осуществления поиска информации, редуцированной по условиям запроса. В этом разделе базы данных предусмотрена возможность, во-первых, поиска по содержанию заголовков дел из архивных описей, введенных в базу данных, и, во-вторых, по всему массиву каталога, в котором, наряду с заголовками дел, содержатся отдельные записи целевой (плановой) или попутной каталогизации, а также локальные блоки информации, которые были конвертированы в базу данных из ранее созданных в Архиве РАН тематических информационных массивов

Пользование электронным каталогом не предусматривает администрирования, поисковые образы не привязаны к классификаторам, словарям или рубрикаторам. Запрос в разделе «Каталог» базы данных осуществляется по ключевым словам в ограничиваемом только здравым смыслом количестве. Работа с интерфейсом «Поиск по каталогу» не требует специальной подготовки, он привычен и удобен любому пользователю, знакомому с поиском в сети Internet.

Кроме того, в Архиве РАН разработан программный модуль «Тайпер», предназначенный для набора заголовков дел описей и каталожных карточек на отдельные дела и документы с последующим экспортом данных в базу данных «Учет и каталог».

В 2011 г. разработана «Информационная система Архива РАН», которая позволяет пользователям в режиме удаленного доступа в сети Internet на сайте isaran.ru ознакомиться со списком фондов, реестром описей и каталогом к документам Архива РАН.

Архивная опись является главным архивным справочником, в котором содержится информация о составе и содержании документов фондов. Одновременно архивная опись - обязательный учетный документ архива, предназначенный для поединичного и суммарного учета документов, закрепления порядка их систематизации, учета изменений в составе и объеме документов. Традиционно, архивная опись - служебный документ, который не подлежит копированию без специального разрешения администрации архива. Разумный баланс в использовании архивных описей, как необходимого справочника в любой исследовательской и архивной работе, с одной стороны, и документа ограниченного использования, не предназначенного для тиражирования, с другой, - залог эффективной работы Архива РАН как учреждения, обладающего значительным информационным потенциалом.

Для реализации возможности изучения пользователями архивных описей архива через Информационную систему Архива РАН была поставлена задача создания базы данных «Электронные описи Архива РАН», которая была выполнена при финансовой поддержке Российского гуманитарного научного фонда в рамках проекта № 11-01-12045в. В течение 2011-2013 гг. предполагалось поэтапно выполнить работу по

оцифровке 2 632 описей фондов Архива РАН.

К реализации проекта были привлечены высококвалифицированные специалисты - старшие научные сотрудники и старшие лаборанты (с высшим образованием) Архива РАН с многолетним опытом работы с архивными документами, хорошо ориентирующиеся в составе документов архива, обладающие знаниями по истории науки и Российской академии наук, владеющие навыками работы на персональных компьютерах и документальных сканерах. Техническая часть исполнения проекта была возложена на штатного инженера для осуществления контроля за состоянием компьютерной техники, сканеров, серверов, сетей, соблюдением политики безопасности работы с данными, модернизацией оборудования и совершенствованием программного обеспечения.

В 2011 г. были проанализированы научно-технологические процессы, связанные с созданием базы данных «Электронные описи Архива РАН». Проведены подбор, расшивка, сканирование 890 описей 425 личных фондов ученых (академиков и членов-корреспондентов Российской академии наук, выдающихся ученых прошлого и современности), проверка результатов сканирования, брошюровка описей. По мере подготовки графических файлов с оцифрованными листами традиционных описей фондов проводилась сверка электронной версии описи с первоисточником.

После выявления и устранения технических погрешностей, возникших при оцифровке, массив с графическими образами описей фондов был конвертирован в специальный раздел базы данных «Учет и каталог» информационной системы «Архив РАН». Результатом заполнения поля с метаданными «Изображения» в разделе «Описи» в распределенной структуре базы данных «Учет и каталог» у сотрудников и исследователей читального зала архива появились принципиально новые возможности доступа к информации - осуществлять просмотр и чтение электронных копий описей в локальной сети архива в режиме реального времени. Одновременно массив с

графическими образами описей фондов размещен в разделе «Реестр описей» на сайте isaran.ru (в режиме «превью»). В 2012 г. отсканированы 1 632 описи 405 фондов учреждений и организаций РАН. Одновременно пользователи сайта получили возможность просматривать и изучать описи в режиме реального времени.

Основной проблемой, возникшей при распознавании отсканированного текста описей, являлось отсутствие в заголовках дел названия фондообразователя. В ручном режиме проводилась работа по преобразованию отсканированных текстов в необходимый шаблон, которые затем с помощью специально разработанного конвертора были загружены в программный модуль «Тайпер». В программном модуле проводилось полное редактирование данных перед конвертацией их в электронный каталог информационной системы «Архив РАН».

В 2013 г. проводится комплекс работ по сканированию описей разрядов (коллекций) Архива РАН, а также дополнений к описям фондов и фондов, поступивших в архив в 2011-2013 гг. В рамках пилотного проекта, включающего распознавание текста оцифрованных описей, корректировку структуры документа и загрузку в конвертер, начато преобразование оцифрованных описей и распознанных текстов в специальный раздел базы данных «Учет и каталог» и размещение на сайте isaran.ru. Осуществление пилотного проекта по распознаванию текстов оцифрованных описей показало, что в автоматизированном режиме этот процесс наиболее эффективен для описей фондов учреждений и организаций РАН. Опыт работы по корректировке структуры заголовков дел описей личных фондов ученых показал, что проведение этого комплекса работ наиболее рационально в ручном режиме в связи с его трудоемкостью.

Перспективными направлениями развития базы данных «Электронные описи Архива РАН» являются: наполнение базы данных «Учет и каталог» как оцифрованными описями, так и заголовками дел описей фондов Архива РАН; внедрение программного

обеспечения информационной системы «Архив РАН» в архивах РАН с постоянным и постоянно-переменным составом документов РАН для подготовки электронных описей в рамках создания Центрального фондового каталога РАН.

Создание базы данных «Электронные описи Архива РАН» позволит исследователям читального зала Архива РАН и пользователям сети Интернет просматривать и изучать описи в режиме реального времени. Одновременно расширилась источниковедческая база для поиска информации в разделе «Учет и каталог» информационной системы Архива РАН.

Список литературы

Ларин М.В. Управление документацией в организациях. М.: Научная книга, 2002. 286 с.

Ларин М.В., Рысков О.И. Электронные документы в управлении. Научно-методическое пособие. М., 2008. 208 с.

Ларин М.В. О государственной политике США по управлению официальными документами. Доклад. XIX Международная научно-практическая конференция «Документация в информационном обществе: «облачные» технологии и электронный документооборот». Федеральное архивное агентство, Всероссийский научно-исследовательский институт документоведения и архивного дела. Москва, 24-25 октября 2012 г.

Ларин М.В. Управление электронными документами. Зарубежная практика // Делопроизводство. 2013. № 1 [Электронный ресурс]. Режим доступа: URL: http://www.top-personal.ru/officeworkissue.html?277

Moore R.W. Building Preservation Environments with Data Grid Technology. AA. 2006. Vol.69. № 1. P.139-158.

Underwood W., Hayslett M., Isbell Sh. и др. Advanced Decision Support for Archival Processing of Presidential Electronic Records: Final Scientific and Technical Report, September 22, 2006 – September 21, 2009. Atlanta, 2009. 37 pp.

Managing Government Records. Memoramdum for the Heads of Executive Departments and Agencies. November 28, 2011 [Электронный ресурс]. Режим доступа: Office of the Press Secretary The White House URL: http://www.whitehouse.gov

You can read completely article in the russian historic-archival magazine “The Herald of an Archivist”. Read more about terms of subscription here .

Полностью материал публикуется в российском историко-архивоведческом журнале ВЕСТНИК АРХИВИСТА. Ознакомьтесь с условиями подписки здесь .