

Электронные документы научно-технической информации в системе ВНТИЦ

© Ю.И. Голосов, Г.А. Брагина, М.Н. Пржиялковская

Всероссийский научно-технический информационный центр (ВНТИЦ)

postmaster@vntic.org.ru

Аннотация

Данное сообщение посвящено системе приема непубликуемых документов научно-технической информации как составляющей системы документооборота Всероссийского научно-технического информационного центра (ВНТИЦ), предназначенного для создания и поддержки интегрального электронного автоматизированного хранилища ВНТИЦ, с предоставлением онлайн-доступа через сеть Интернет к его ресурсам.

Всероссийский научно-технический информационный центр (ВНТИЦ) формирует Федеральный фонд по непубликуемым источникам информации – отчетам о научно-исследовательских и опытно-конструкторских работах, кандидатским и докторским диссертациям, описаниям алгоритмов и программ, осуществляет ведение Единого реестра результатов научно-технической деятельности.

В настоящее время фонд ВНТИЦ насчитывает более 7 миллионов документов. Ежегодные поступления во ВНТИЦ составляют около 100 тысяч различных документов, отражающих контент научно-технической информации (НТИ). Обработка такого количества документов требует значительных финансовых и временных затрат.

Поставщиками информации в коллекцию являются организации науки и высшей школы, промышленные предприятия – исполнители НИОКР, диссертационные советы и авторы диссертаций, а также бюджетополучатели – министерства и ведомства, выступающие государственными заказчиками НИОКР.

Любой ресурс, в том числе и информационный, становится бесполезным и не может быть использован без инструментов систематизированного хранения, доступа и распространения.

Современное состояние ИТ-отрасли позволяет создавать информационные системы любой степени сложности, в том числе и основанные на принципах электронного документооборота. Однако, основная

масса разрабатываемых систем электронного документооборота, в том числе и ориентированных на использование и ведение депозитариев научно-исследовательских проектов и разработок, рассчитаны на использование в одной организации с определенной предметной областью и определенным типом информации, а также ограниченным количеством пользователей.

Система электронного документооборота научно-технической информации в федеральном информационном центре имеет ряд принципиальных особенностей. Прежде всего, это система, рассчитанная на прием, обработку, хранение и распространение больших объемов информации – несколько сотен тысяч документов в год. При этом объем документов колеблется от 3 Кб до 250 Мб. Информация, представленная в документах, не структурирована либо слабо структурирована, документы относятся к различным областям знаний, т.е. фонд политематичен.

Функционально система электронного документооборота научно-технической информации в Федеральном информационном центре должна решать следующие задачи:

1. Подготовка и прием электронных форм документов в интерактивном режиме;
2. Смысловая обработка документов (индексирование, рубрицирование, редактирование и т.д.);
3. Организация электронного хранилища данных с возможностью онлайн-доступа пользователей;
4. Выпуск электронных изданий;
5. Прием и выполнение заказов на копии первоисточников;
6. Аналитическая обработка и статистический анализ контента фонда;
7. Страхование копирования и хранение документов.

Традиционно, документы, представляемые во ВНТИЦ, поступали только на бумаге и, пройдя определенные стадии обработки, трансформировались в электронный вид. По мере развития компьютерных технологий и внедрения их в технологический процесс формирования федерального фонда по непубликуемым источникам информации, различные операции по обработке, хранению и распространению информации автоматизировались и модернизировались, создавая основу для системы электронного документооборота.

Труды 10-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» – RCDL'2008, Дубна, Россия, 2008.

Было создано интегральное электронное автоматизированное хранилище ВНТИЦ, включающее банк данных государственных контрактов на НИОКР, политематические ретроспективные реферативно-библиографические базы данных по государственной регистрации и учёту НИОКР и диссертаций, а также хранилище полнотекстовых отчётов и диссертаций с реализацией организации онлайн-доступа пользователей через сеть Интернет к ресурсам электронного хранилища. Однако, операции по приему документов до настоящего времени оставались неизменны.

Прием электронных форм документов в сетевом режиме имеет ряд особенностей:

1. Прежде всего, в существующей практике юридическую силу по-прежнему имеют документы на бумажном носителе, поэтому наряду с электронной формой документа необходимо иметь его подтверждение на бумаге, оформленное по всем правилам, предусмотренным для такого рода документов.

2. В Российской Федерации существуют документы, регламентирующие бумажное представление документов научно-технической информации (ГОСТ 7.32-2001 от 1.07.2002 – Отчет о научно-исследовательской работе, структура и правила оформления), но отсутствуют стандарты или рекомендации, регламентирующие электронное представление как текстов документов научно-технической информации, так и их нетекстовых частей (формул, графических элементов, фотографий и т.д.).

3. Значительный объем полнотекстового документа (отчёта о НИОКР, диссертации) – до 250 Мб.

В настоящее время во ВНТИЦ разработана первая очередь системы электронного документооборота НТИ, включающая как технологию и средства приема, так и обработки, хранения и распространения реферативной информации.

Прием документов в электронном виде (регистрационных, информационных карт НИОКР, информационных карт диссертаций) реализован с использованием технологии ASP.NET. Доступ к системе осуществляется через сайт ВНТИЦ – www.vntic.org.ru. Далее принятые документы поступают в технологическую базу, обрабатываются и загружаются в электронное автоматизированное хранилище ВНТИЦ. Первоисточники – диссертации, тексты НИР и ОКР – сканируются и также поступают в электронное хранилище.

В 2009 году планируется проведение разработки второй очереди системы документооборота, ориентированной на работу с полнотекстовыми документами.

Литература

- [1] ГОСТ 7.32-2001 Межгосударственный стандарт. Отчет о научно-исследовательской работе (Структура и правила оформления).
- [2] Богданов Ю.М., Павлов Л.П. и др. Отчет НИОКР «Развитие информационной системы и

технологий формирования, ведения и распространения информации Единого реестра результатов научно-технической деятельности и государственного фонда неопубликованных документов», 2007 г., с 436, ВНТИЦ.

- [3] Л.Г. Титарев, Ю.М. Богданов и Л.П. Павлов “Digital Documents in Grey Literature : New Challenges ”, доклад, с. 7, Антверпен (Бельгия), 9-я Международная конференция по серой литературе GL 9.

Electronic Documents of Scientific and Technical Information in VNTIC System

Y.I. Golosov, G.A. Bragina, M.N. Przyjalkovskaya

The report describes the system of receiving of scientific and technical information documents, as the part of The Scientific and Technical Information Centre of Russia (the Russian acronym – VNTIC) documents circulation system, intended to provide the generation and support of VNTIC integral electronic documents repository, provided with Internet online information search.