

Средства резервного копирования и восстановления коллекций информационных ресурсов в системе Соционет¹

М.Р. Когаловский
Институт проблем рынка РАН

С.И. Паринов
Центральный экономико-математический институт РАН

Аннотация

В системе Соционет в настоящее время накоплен большой объем информационных ресурсов. Одним из основных источников для их пополнения и актуализации являются автономно создаваемые в среде системы Соционет с помощью ее сервисов наборы коллекций метаданных. Их администраторы помимо доступа к указанным сервисам получают при регистрации в системе пространство системной памяти для создания и поддержки метаданных, их предоставления для публичного доступа пользователям системы. Для обеспечения надежности поддержки таких автономно поддерживаемых информационных ресурсов необходимы специальные средства. В статье обсуждаются функциональные возможности предназначенных для этих целей сервисов системы Соционет, обеспечивающих резервное копирование и восстановление таких автономно поддерживаемых коллекций метаданных.

Источники информационных ресурсов Соционет

Информационные ресурсы Соционет (метаданные, описывающие информационные объекты, доступ к которым предоставляет пользователям система) пополняются из нескольких источников. Одну их группу составляют открытые архивы, зарегистрированные в системе. Сборщик метаданных Соционет, реализующий протокол OAI-PMH [1] технологии открытых архивов (Open Archives Initiative, OAI) [2], в соответствии с установленным регламентом анализирует их репозитории метаданных и в соответствии с их изменениями актуализирует и пополняет репозиторий метаданных Соционет.

Другая группа источников – это личные метаданные, автономно создаваемые сотрудниками институтов РАН и других организаций либо уполномоченным сотрудником организации от их имени. Лица, ответственные за указанную деятельность, называются администраторами соответствующих коллекций. Указанные метаданные организуются в виде материализованных коллекций, которые фактически являются представителями виртуальных коллекции описываемых ими информационных объектов. Эти коллекции метаданных автономно создаются их администраторами, актуализируются, а также сохраняются с использованием ресурсов и сервисов системы.

¹ Работа поддержана РГНФ, проект 11-02-12026-в

Личная зона администратора коллекций метаданных

Для зарегистрировавшихся в системе и получивших полномочия от администратора Соционет администраторов коллекций метаданных, система предоставляет комплекс системных сервисов, обеспечивающий их деятельность, а также ресурсы пространства памяти для создания, актуализации и автономного хранения коллекций метаданных. Выделенная администратору область пространства памяти Соционет для автономной его работы с управляемыми им коллекциями называется *личной зоной администратора*.

Как уже отмечалось, информационные ресурсы в личной зоне администратора представлены в виде совокупности коллекций метаданных, описывающих информационные объекты различных типов, поддерживаемых в системе Соционет. Эти коллекции метаданных доступны только администратору, учредившему эту личную зону, и таким образом представляют собой его автономный ресурс. В следующем разделе приводится более точное обсуждение характера, смыслового содержания и представления информационных ресурсов, создаваемых и поддерживаемых администратором в личной зоне.

Информационные ресурсы в личной зоне администратора

Что же представляют собой информационные ресурсы, создаваемые и поддерживаемые администраторами в их личных зонах? Напомним, что система Соционет основана на технологии *Инициативы открытых архивов* [2]. В соответствии с этой технологией, информационные ресурсы, непосредственно поддерживаемые системой, организованы в виде репозитория метаданных, описывающих свойства и адреса в среде Веб тех информационных объектов (статей, монографий, диссертаций, научных отчетов и т.д.), которые представляют или могут представлять интерес для пользователей системы. Поддержка самих этих информационных объектов не входит в задачу системы.

Когда администратор информационных ресурсов создает в своей личной зоне некоторую коллекцию, он фактически создает коллекцию метаобъектов (описателей) информационных объектов одного из допустимых в системе типов. Именно такие коллекции метаобъектов составляют содержимое личной зоны этого администратора. Как указывалось выше, эти коллекции фактически воспринимаются их создателями и пользователями системы как коллекции описываемых ими информационных объектов или иначе как виртуальные коллекции информационных объектов.

Каждый из рассматриваемых метаобъектов порождается в личной зоне при заполнении администратором специальной экранной формы. В ее окна вписываются идентификатор описываемого информационного объекта, уникальный в информационном пространстве системы, и ряд его атрибутов, состав которых зависит от его типа. Когда заполнение формы закончено, системные механизмы генерируют на основе ее содержимого представление данного описателя в виде XML-документа, который сохраняется в личной зоне как один из элементов соответствующей коллекции метаданных.

Таким образом, коллекция информационных объектов представляется в личной зоне системы Соционет как коллекция XML-документов, являющихся описателями соответствующих информационных объектов. В каждом описателе – XML-документе -

содержится гиперссылка на сам описываемый информационный объект, который может храниться на любом сайте в Интернет или в любом FTP-архиве.

Соционет является научной системой. Наряду с коллекциями описателей научных информационных объектов, примеры которых приведены выше, в ее репозитории метаданных представлены также коллекции профилей (описаний) научных и образовательных организаций, а также профилей их сотрудников. Такие коллекции могут также формироваться и поддерживаться автономно в личных зонах администраторов.

Трансляция коллекций личной зоны

Администратор коллекций метаданных обладает полномочиями на предоставление открытого доступа к коллекциям описываемых ими информационных объектов пользователям системы Соционет. Для этого он имеет возможность включить в соответствии с технологией Открытых архивов режим трансляции той или иной своей коллекции в зону публичного доступа пользователей (в так называемую *публичную зону*). В результате трансляции коллекции ее копия создается в публичном (общедоступном) информационном пространстве Соционет, и состояние этой копии в соответствии с установленным регламентом синхронизируется с состоянием оригинала этой коллекции в личной зоне. Состояние коллекции в личной зоне может изменяться администратором путем пополнения коллекции, обновления уже существующих в ней описателей или их удаления.

Администратор коллекции может также отключать трансляцию этой конкретной коллекции. При этом соответствующая коллекция информационных объектов станет невидимой для пользователей системы в публичной зоне.

Результаты действий в их личных зонах администраторов коллекций, предоставляемые для трансляции в публичную зону, модерируются администратором Соционет и становятся видимыми пользователям в соответствии с регламентом, установленным в системе. В настоящее время в Соционет это происходит на следующий день после осуществления операций администратором в его личной зоне.

Резервная копия коллекций личной зоны

В настоящее время в системе Соционет зарегистрировано несколько десятков администраторов информационных ресурсов, поставляющих с использованием личных зон ресурсы в Соционет для публичного доступа. В их личных зонах накоплено значительное количество коллекций. Многие из них являются довольно крупными.

В этих условиях злободневной стала проблема создания в Соционет сервисов, которые бы позволяли создавать резервные копии (backup copy) содержимого личных зон и восстанавливать их содержимое в случае разрушения с помощью этих резервных копий. Такие сервисы были разработаны.

Для доступа к сервису создания резервной копии контента личной зоны необходимо после входа в личную зону с использованием идентификатора пользователя (логина) и пароля выбрать в меню, предлагаемом в нижних строках экрана начальной страницы личной зоны режим «Управление коллекциями» (рис. 1).

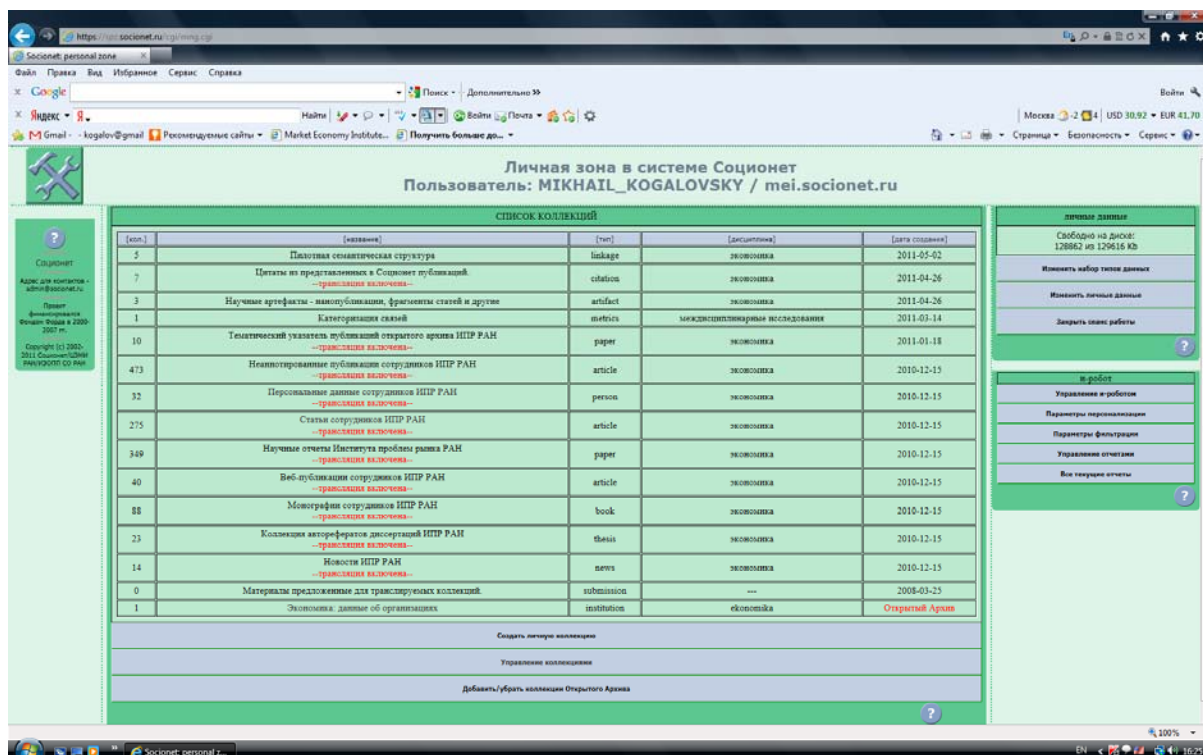


Рис. 1. Начальная страница личной зоны

На появившемся экране режима управления коллекциями (рис. 2) в меню, показанном в верхней его части, нужно выбрать пункт создания резервной копии и восстановления («Бэкап и восстановление»). В результате появится экран, соответствующий этим функциям (рис. 3).

Для создания резервной копии нужно, прежде всего, пометить флажки в левом столбце списка коллекций, содержащихся в личной зоне, тех коллекций, которые необходимо включить в состав резервной копии. Виртуальная кнопка в левом верхнем углу экрана позволяет одновременно пометить все эти коллекции. Завершается процесс создания резервной копии кликом мышью на виртуальной кнопке «Выполнить» в левом нижнем углу экрана. Система сохраняет сформированную резервную копию в указанном администратором месте на диске. Файлу копии система присваивает некоторое сгенерированное имя. Однако администратор при желании может переименовать этот файл.

Резервная копия представляет собой ZIP-архив всех коллекций XML-документов – описателей информационных объектов, которые составляют данные коллекции. Эти XML-документы являются небольшими текстовыми файлами. Так как текстовые файлы существенно сжимаются при включении их в ZIP-архив и, к тому же, в исходном виде имеют небольшой объем, резервная копия имеет относительно небольшой объем даже при наличии довольно крупных по количеству описателей коллекций в личной зоне.

Сервис системы Соционет для создания резервных копий контента личной зоны наряду с обеспечением возможности восстановления его состояния в случае разрушения может быть использован и для другой цели. В частности, его можно использовать в системе

Соционет для обмена коллекциями между различными личными зонами. Каким образом можно осуществить такую возможность, рассматривается ниже.

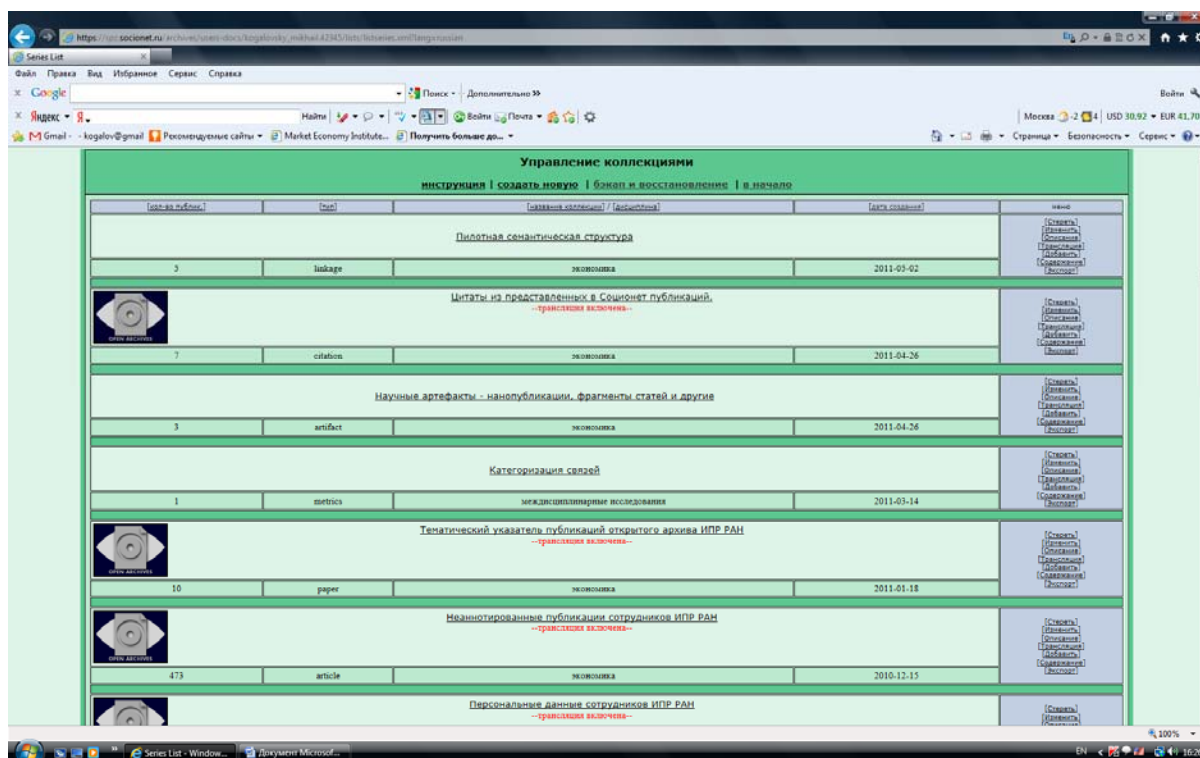


Рис. 2. Экран режима управления коллекциями

Восстановление личной зоны

В случае разрушения контента личной зоны некоторого администратора информационных ресурсов можно восстановить полный контент этой личной зоны или отдельных ее коллекций на момент создания последней резервной копии. Можно также восстановить состояние коллекций личной зоны на момент создания какой-либо иной резервной копии.

Для решения указанных задач необходимо поступить следующим образом. Прежде всего, следует описанным выше способом войти в личную зону и выполнить действия, которые приведут к появлению экрана создания резервной копии и восстановления (рис. 3). Далее следует пометить флажки тех коллекций в списке коллекций, состояние которых нужно восстановить. После этого нужно в верхней части экрана в меню, размещенном под его заголовком, выбрать функцию восстановления («Восстановить из бэкап-файла») и по запросу системы указать местонахождение и имя ZIP-файла резервной копии, которую необходимо использовать для восстановления. Таким образом, задача будет решена.

Другое использование сервиса создания резервной копии/восстановления

Сервис системы Соционет, позволяющий создавать резервные копии всего множества коллекций личной зоны или некоторого его подмножества может использоваться и для другой цели – для обеспечения обмена коллекциями между личными зонами системы. По договоренности администраторов этих личных зон создается резервная копия

экспортируемых коллекций исходной личной зоны. Эта резервная копия может использоваться далее с помощью функции восстановления для импорта требуемых коллекций в целевую личную зону. Естественно, операции создания резервной копии и включения экспортируемых коллекций в целевую личную зону могут выполняться только их администраторами. Таким образом, в данном случае сервис восстановления выполняет функцию экспорта коллекций.

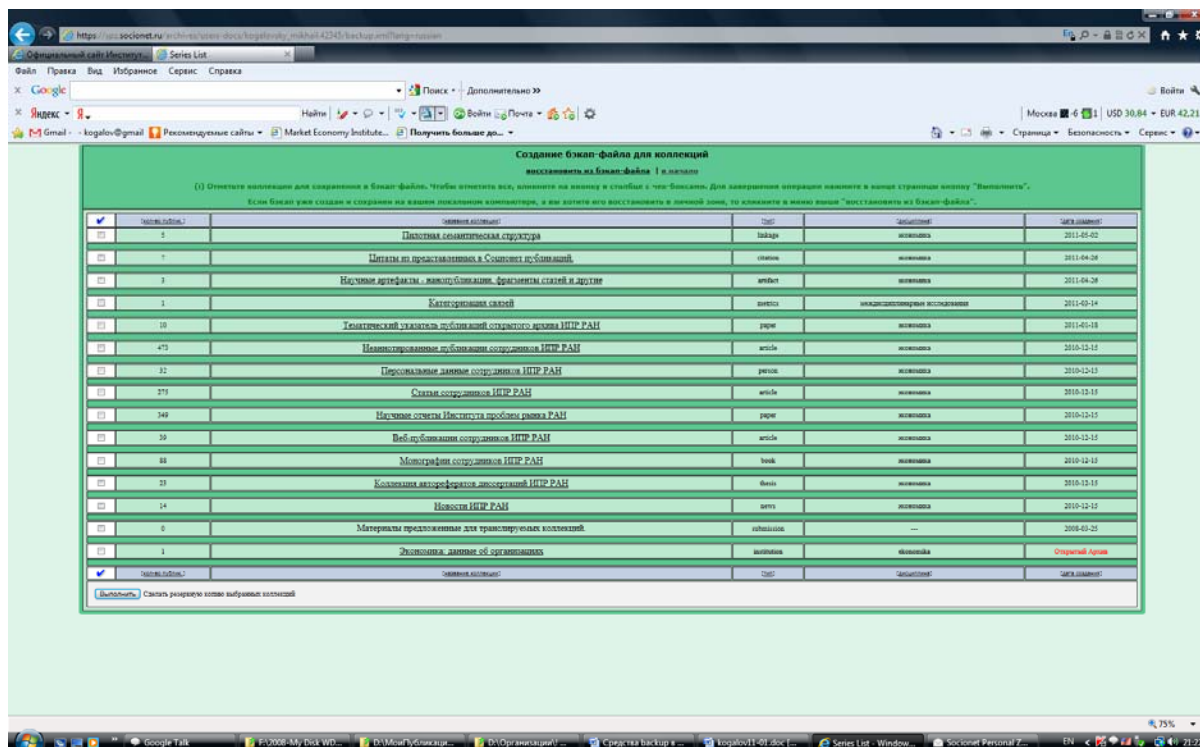


Рис. 3. Экран создания резервной копии и восстановления контента личной зоны

Заключение

Более подробную информацию об организации информационных ресурсов системы Соционет можно найти, например, в работах [3, 4].

Литература

1. The Open Archives Initiative Protocol for Metadata Harvesting. Protocol Version 2.0 of 2002-06-14. <http://www.openarchives.org/OAI/openarchivesprotocol.html>
2. Open Archives Initiative. <http://www.openarchives.org/>
3. Когаловский М.Р., Паринов С.И. Информационные ресурсы, наукометрические показатели и показатели качества метаданных системы Соционет. Труды 9-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции - RCDL'2007, Переяславль-Залесский, Россия, 15-18 октября 2007 г. - Переяславль-Залесский: Изд-во «Университет города Переяславля», 2007. – С. 45-54.
4. Когаловский М.Р., Паринов С.И. Метрики онлайн-информационных пространств // Экономика и математические методы. – 2008. – Вып. 2