

Корпоративный каталог СО РАН

© Жижимов О.Л.,
ИГМ СО РАН
Новосибирск
zhizhim@nsc.ru

© Турпанов А.А.,
ИКФИА СО РАН
Якутск
turpanov@ikfia.ysn.ru

© А.М. Федотов
ИВТ СО РАН
Новосибирск
fedotov@sbras.ru

Аннотация

В докладе рассматриваются принципиальные вопросы построения распределенной информационно-вычислительной системы Сибирского отделения и корпоративного LDAP-каталога СО РАН как ее основы. Обсуждается структура каталога и его содержание. Отдельно рассматриваются административные мероприятия по построению этого каталога.

1 Проблемы интеграции ресурсов

Проблема доступа к информационным и вычислительным ресурсам является одной из основных проблем, возникающих в деятельности научного сообщества. В настоящее время наблюдаются переход к исключительно распределенной схеме создания, поддержания, хранения ресурсов и, в то же время, – стремление к виртуальному единству посредством предоставления свободного доступа к любым ресурсам сети через ограниченное число «точек доступа». Осознание необходимости интеграции разнородных научных ресурсов привело к созданию интегрированных (единых) *научных информационных систем* (НИС), которые позволили бы установить связи между разнородными документами, организовывать единые каталоги документов, а также создавать специализированные системы поиска.

Основная проблема, связанная с функционированием интегрированных территориально распределенных информационных систем (как ИРИС в СО РАН [1], так и ЕНИП в РАН [2-3]), широко известна – это практически реально не функционирующая система актуализации информации. Как показал наш опыт, решить эту проблему административными методами практически невозможно. Отметим, что эффективная эксплуатация информационных ресурсов возможна только в том случае, когда они *постоянно поддерживаются авторами*.

По-видимому, единственный путь к решению этой проблемы состоит в интеграции в рамках интегрированной распределенной информационной системы данных локальных информационно-

справочных систем, существующих в Институтах, и придания этим системам функций глобальной (корпоративной) аутентификации и авторизации пользователей по доступу к информационным ресурсам.

Интеграция сетевых сервисов и информационно-вычислительных ресурсов в *корпоративных распределенных информационных системах* (КРИС) с множеством управляющих серверов требует создания единой *корпоративной распределенной справочной системы* (КРСС) о конфигурации и содержании собственно КРИС. В основу такой системы может быть положена уже хорошо зарекомендовавшая себя в сети служба каталогов, функционирующая в соответствии рекомендациями ITU X.500. В этом смысле интегрированная информационная система Отделения (а может быть и РАН в целом) должна представлять собой единый (возможно распределенный) каталог всех ресурсов, созданный в соответствии с международными рекомендациями (в целях интероперабельности). Данный подход уже хорошо зарекомендовал себя в области построения распределенных каталогов обычных библиотек

Основными предпосылками внедрения КРСС являются:

- Необходимость учета и поддержания в актуальном состоянии информации о сервисах и ресурсах КРИС;
- Необходимость организации единой политики доступа к ресурсам КРИС на основе единых правил;
- Необходимость единого учета пользователей КРИС;
- Необходимость однократной авторизации пользователей КРИС;
- Необходимость децентрализованного формирования специальных справочников и обеспечения доступа к ним через различные интерфейсы. Эти справочники (созданные в соответствии с международными рекомендациями по структуре мета-данных) должны содержать:
 - Информацию о доступных в КРИС сетевых сервисах (как основных FTP, SMTP, WWW и т.п., так и дополнительных) с адресами обслуживающих их серверов;
 - Информацию о доступных в КРИС информационных и вычислительных ресурсах с указанием точки и метода доступа к ним и используемых схем данных;
 - Информацию об администраторах КРИС;

Труды 8^{ой} Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2006, Суздаль, Россия, 2006.

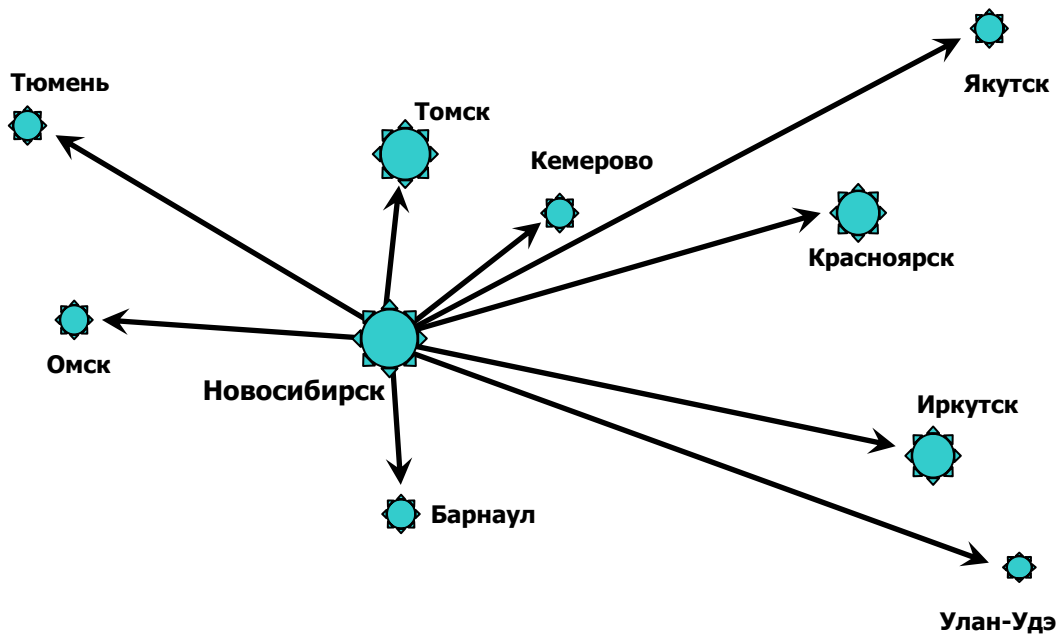


Рис.1 Научные центры СО РАН

- Метаданные об организациях, участвующих в использовании, создании и поддержке КРИС;
- Метаданные о персонах (включая контактную информацию);
- И т.д. и т.п.

2 Основные требования к КРСС

Сформулируем основные требования, которым должна подчиняться корпоративная распределенная справочная система:

- Организация информации в КРСС должна обеспечивать ее сегментацию на отдельные административные сегменты.
- Актуальность информации в КРСС должна поддерживаться множеством администраторов независимо в каждом сегменте системы.
- Доступ к КРСС должен быть основан на открытых стандартах (OSI).
- КРСС должна обеспечивать возможность хранения информации различного типа с возможностью ее поиска по различным атрибутам.

На сегодняшний день наиболее распространенной технологией создания КРСС является технология LDAP. Для создания КРСС на основе технологий LDAP в СО РАН в настоящее время проводится ряд мероприятий в рамках корпоративных информационных проектов. Эти работы призваны обеспечить существующие информационные системы, в том числе и распределенные, созданные на основе технологий Z39.50 и WWW, единой КРСС, представленной корпоративным каталогом LDAP.

Корпоративный каталог СО РАН создается как централизованно-распределенная информационная система с разграничением зон ответственности за актуальность информации между центрами и организациями СО РАН. Территориальный охват ката-

лога (см. рис.1) – все научные центры СО РАН (Новосибирск, Томск, Омск, Красноярск, Иркутск, Кемерово, Якутск, Тюмень, Улан-Удэ). При этом структура каталога отражает административную структуру СО РАН, а минимально возможной административной единицей каталога, административным доменом, является организация СО РАН (см. рис. 2).

3 Структура каталога

Ветка дерева каталога каждой организации (см. рис. 3) включает на первом этапе следующие типы информации (в дальнейшем планируется расширение):

- **Информация о научных центрах** содержится в объектах класса *sbrasOu*, определяющего набор атрибутов (название, реквизиты, URL и т.п.), характеризующих научный центр. Объекты этого класса присутствуют на первом уровне дерева каталога СО РАН.
- **Информация об организации** содержится в объекте класса *sbrasOrg*, определяющего набор атрибутов (название, реквизиты, URL и т.п.), характеризующих организацию. Объекты этого класса могут присутствовать в узле-контейнере класса *sbrasOu* (научный центр) или в узле класса *sbrasOrg* (объединенный институт).
- **Информация о персонах** содержится в объектах класса *sbrasPerson*, определяющего набор атрибутов (фамилия, имя, телефоны, должность и т.п.), характеризующих персону. Объекты этого класса группируются в узле-контейнере с именем *People* класса *OpenLDAPou*.
- **Информация о сетевых сервисах**, которые организация предоставляет для внешнего доступа,

содержится в объектах класса, наследующего свойства стандартных классов *ipProtocol*, *ipHost* и *ipService*. Объекты этого класса группируются в узле-контейнере с именем *NetServices*.

- **Информация об информационных ресурсах**, к которым организация предоставляет доступ, содержится в объектах класса *dcContainer*, опреде-

- **Информация об адресном пространстве IP**, которое занимает соответствующая организационная структура СО РАН, содержится в объектах стандартного класса *ipNetworkNumber*. Объекты этого класса группируются в узле-контейнере с именем *networks* стандартного класса *organizationalUnit*.

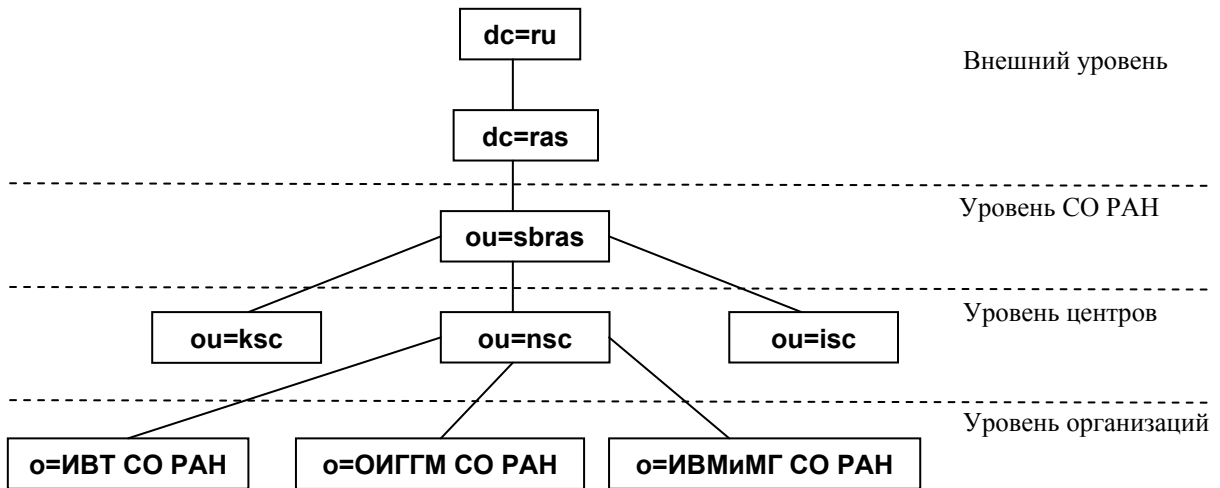


Рис.2 Фрагмент LDAP-дерева каталога СО РАН

ляющего атрибуты в соответствии со схемой *Dublin Core*. Объекты этого класса группируются в узле-контейнере с именем *InfResources*.

- **Другая информация**, необходимая для соответствующей организации, но не регламентируемая корпоративными документами.

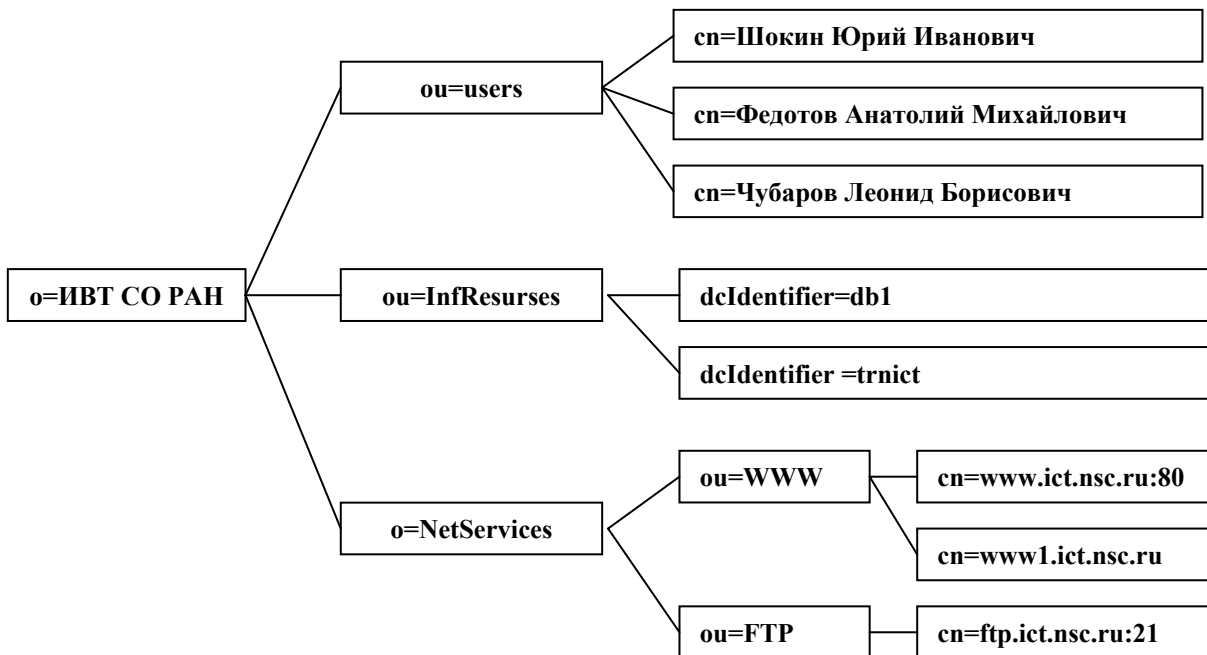


Рис.3 Фрагмент каталога организации

Наличие в корпоративном каталоге вышеуказанной информации в соответствующих схемах позволяет различным сетевым приложениям и сервисам ее использовать. Среди таких приложений можно выделить

- PROXY-сервисы, ограничивающие обслуживание списками пользователей и компьютеров;

- WEB-серверы, регламентирующие доступ к WEB-сайтам в соответствии с LDAP-сертификатом;
- Вычислительные серверы, регламентирующие доступ к вычислительным сайтам в соответствии с LDAP-сертификатом;
- Почтовые серверы, использующие списки рас-

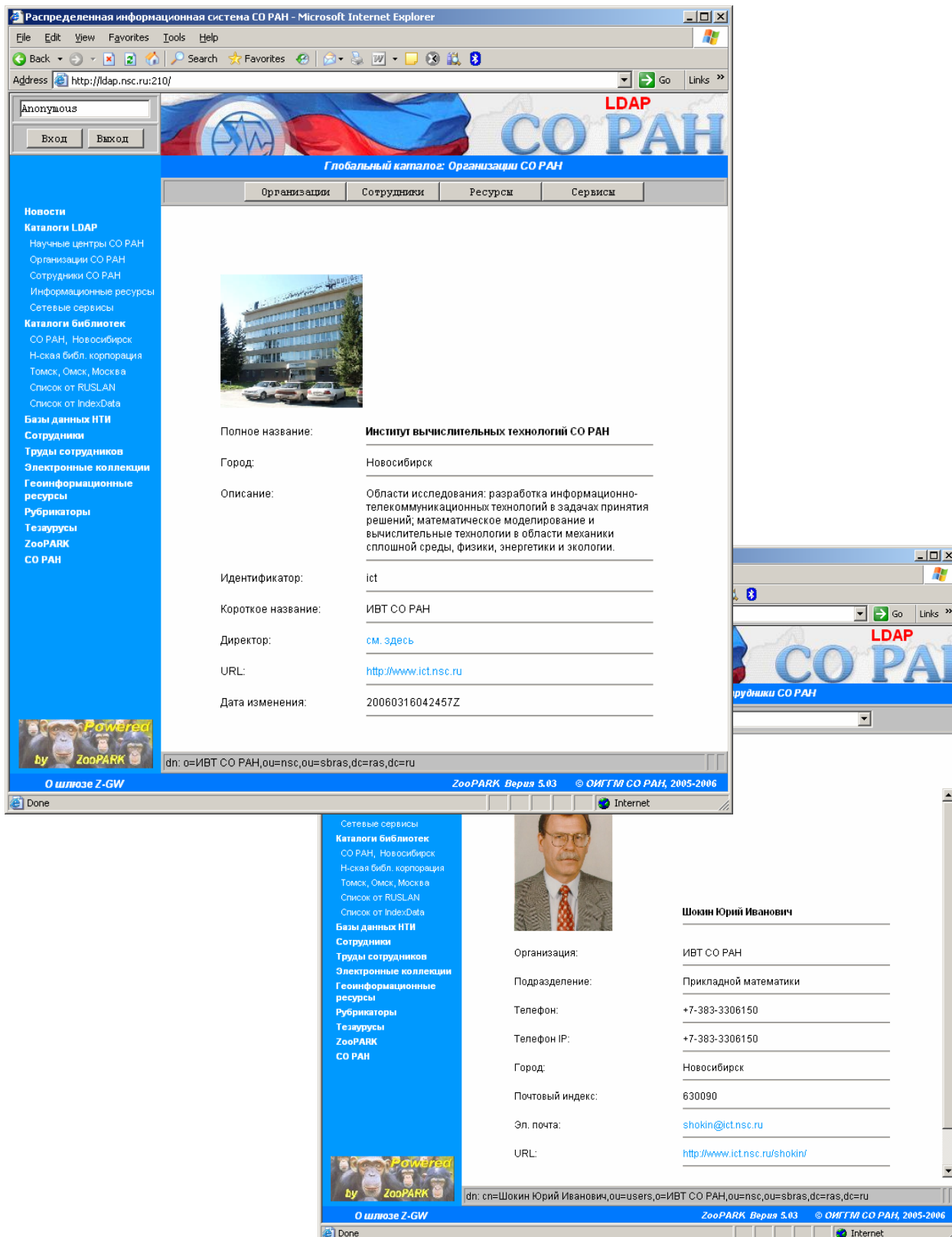


Рис.4 Шлюз WWW-LDAP (WWW-Z39.50) информационной системы СО РАН

ссылки по группам пользователей каталога LDAP;

- Клиенты почтовых служб, использующие информацию о персонах для формирования адресных книг;
- Серверы баз данных, в том числе и серверы Z39.50, регламентирующие доступ к информационным ресурсам по именам пользователей, LDAP-сертификатам или IP-адресам клиентов;
- Сервисы учета, используемых вычислительных, информационных и телекоммуникационных ресурсов
- Другие сервисы и приложения.

Корпоративный каталог СО РАН является в первую очередь информационной системой, которая в силу своей организации содержит актуальную информацию. Доступ к части этой информации разрешен для всех пользователей каталога и может предоставляться через специальные шлюзы WWW. В частности, сегодня функционирует шлюз на основе сервера ZooPARK [4] (WWW, Z39.50), интерфейс которого изображены на рисунке 1. (см. также <http://ldap.sbras.ru:210>).

Отдельный интерес представляет последовательность организационных мероприятий, направленных на формирование корпоративного каталога СО РАН и поддержки его в актуальном состоянии. Не секрет, что информационный ресурс будет поддерживаться в актуальном состоянии тогда и только тогда, когда требование актуальности этого ресурса является одним из основных требований нормального функционирования информационного сообщества. Например, предоставление доступа к информационным ресурсам, необходимым для научных исследований, только посредством аутентификации и авторизации пользователей в корпоративном каталоге, приводит к заинтересованности каждого научного сотрудника в отображении в этом каталоге собственной персоны.

На сегодняшний день последовательность организационных мероприятий, направленных на формирование корпоративного каталога, выглядит следующим образом:

- Создание скелета корпоративного каталога на основе фиксированных схем данных, в том числе и собственной разработки.
- Наполнение каталога информацией до уровня организаций.
- Создание WEB-шлюзов для доступа пользователей к корпоративному каталогу как к справочной системе.
- Создание и утверждение нормативной документации (регламента) по ведению корпоративного каталога и разграничению полномочий администраторов.
- Предоставление прав удаленного редактирования информации в глобальном каталоге на центральном сервере представителям научных центров.

- Запрещение доступа к корпоративным информационным ресурсам пользователям, не прошедшим авторизацию в корпоративном каталоге.
- Установка периферийных серверов в научных центрах и организация системы репликаций данных от периферии к центру.

Мы надеемся, что последовательность перечисленных мероприятий позволит создать распределенный корпоративный LDAP-каталог СО РАН, территориально покрывающий более 2/3 территории Российской Федерации.

Литература

- [1] Шокин Ю.И., Федотов А.М., Жижимов О.Л., Мазов Н.А. Интегрированная распределенная информационная система (ИРИС) Сибирского отделения РАН // Выездное заседание координационного научного совета СО РАН по целевой программе: Информационно-телекоммуникационные ресурсы СО РАН, г. Иркутск, 29-30 июля, 2002 г.: Материалы заседания. - Иркутск, СО РАН. Иркутский научный центр. - 2003. - С.139-149.
- [2] Созыкин А.В., Масич Г.Ф., Масич А.Г., Бездушный А. Н. Вопросы интеграции информационных и сетевых служб. Варианты использования LDAP каталогов // Труды 6-ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL2004, Пушкино, Россия, 2004.
- [3] Бездушный А. А., Нестеренко А. К., Сысоев Т.М., Бездушный А. Н., Серебряков В. А. Возможности технологий ИСIP в поддержке Единого Научного Информационного Пространства РАН // Электронные библиотеки. - 7 (6). - 2004.
- [4] Жижимов О.Л., Мазов Н.А. Принципы построения распределенных информационных систем на основе протокола Z39.50. - ОИГГМ СО РАН, Новосибирск: ИВТ СО РАН. - 2004. - ISBN 5-9554-0017-6. - 361 с.

The corporate catalogue of the Siberian Branch of the Russian Academy of Science

Zhizhimov O.L., Turpanov A.A., Fedotov A.M.

Some aspects of construction of the corporate LDAP-catalogue of the Siberian Branch of the Russian Academy of Science are considered. The structure of the catalogue and its contents is discussed. Administrative actions on construction of this catalogue are separately considered.