

Разработка репозитория на базе Dspace: опыт Научно-Медицинской Библиотеки ГУМФ им. Николае Тестемицану

А. Колесников, Л. Малахов

Мотивация

- Изменилась парадигма научной коммуникации и роль публикаций.
- Значительно ускорился обмен данными, включая экспериментальные.
- Архивы препринтов (arxiv.org).
- Не публикуемые формально документы, например, технические отчеты.
- Доступность первичного исследовательского материала в сети → меньшая зависимость от публикаций, возможность проверить и воспроизвести результаты, проанализировать данные.
- Печатные статьи утрачивают функцию научной коммуникации. Сейчас это скорее фиксация результата исследования и гарант научного признания.
- ВУЗы производят сотни тысяч неопубликованных документов, множество технических отчетов.
- Учебные заведения дополнительно производят большое количество первичных цифровых материалов.
- В учебных заведениях производится и распространяется программное обеспечение.
- Проект DSpace отразил новые тенденции в сфере научных коммуникаций и новую роль библиотек в сборе, сохранении и распространении информации.

arXiv.org



Cornell University
Library

arXiv.org

Open access to 1,085,689 e-prints in Physics, Mathematics, Computer Science, Quantitative Biology

Subject search and browse:

10 Sept 2015: [arXiv online fund raising pilot](#)

12 Jan 2015: [Reflections on the 1 million paper milestone](#)

12 Jan 2015: [A project update, including a brief summary of activities in 2014, has been posted](#)

See cumulative "What's New" pages. Read [robots beware](#) before attempting any automated download

Physics

- [Astrophysics \(astro-ph new, recent, find\)](#)
includes: [Astrophysics of Galaxies](#); [Cosmology and Nongalactic Astrophysics](#); [Earth and Planetary Astrophysics](#); [Solar and Stellar Astrophysics](#)
[Condensed Matter \(cond-mat new, recent, find\)](#)

archive.org

INTERNET  ARCHIVE

[ABOUT](#) [CONTACT](#) [BLOG](#) [PROJECTS](#) [♥ DONATE](#) [HELP](#) [TERMS](#) [JOBS](#) [VOLUNTEER](#)

Internet Archive is a non-profit library of millions of free books, movies, software, music, and more.



8.3M



2.1M



2.6M



111K



1.1M



157K



175K

GO

[Advanced Search](#)






Функции

- Загрузка, управление, распространение и описание цифровых материалов
- Интегрированные подсистемы для различных представлений цифровых данных и связанных с ними метаданных
- Гибкие схемы метаданных
- Система индексирования и поиска
- Система управления архивом
- Инструмент для создания политик доступа и управления
- Система генерации постоянных идентификаторов, и пр.



История

- Разработка MIT и HP  
- С 2009 – DuraSpace 
- Руководитель разработки Тим Донохью
- Актуальная версия 5.3 (после неудачной 5.2)
- В конце 2015 будет версия 6, в 2016 – версия 7

Проблемы

- Накоплен опыт
- Структура по подразделениям – это ошибка!
- Половина структурных подразделений исчезла, другая переименовалась а третья добавилась.
- На каждой кафедре есть «публикации сотрудников» и пр. массовые коллекции. Тяжело собрать все публикации вместе на одном экране во одном списке (discover появился недавно); в архиве получается множество разделов и коллекций, и разобраться в них невозможно; практически невозможно использовать маппинг (mapping), перенос и пакетную загрузку
- Постепенно до всех доходит, что структурное подразделение и прочие подобные данные надо засовывать в метаданные. Оказывается, что разделы (community) и коллекции не нужны



Перспективы

- Тим заявил, что в DSpace 7 не будет разделов и коллекций. Сейчас DSpace поддерживает плоские (не-иерархические) схемы метаданных
- Один интерфейс, один поиск, одна статистика
- Web-интерфейс для большинства конфигураций
- Mapping – это возможность нахождения документа в разных коллекциях. Проще сделать через jsru1, особенно, если нет двух коллекций с одинаковым именем.
- Многие отказываются от маппинга, потому что возникают коллизии при экспорте и при работе OAI (Open Archive Interface)



Технологическая цепочка

- Организационная подготовка
- Выбор сервера
- Выбор ОС
- Установка ОС и утилит
- Установка поддержки
- Установка DSpace
- Обслуживание сервера и эксплуатация системы



Организационная подготовка

- Решение на уровне руководителя организации
- Важный вопрос – авторские права
- Выделение персонала
- Выбор некоторых параметров, например, названия системы и URL, включение в DNS
- SMTP сервер
- Получение идентификатора для уникальных ссылок **Handle.Net[®]**



Выбор сервера



- Мощный процессор и большая память не требуются, но надо много места на диске
- Можно разделить на три части – система, база, медиафайлы
- Систему можно ставить на виртуальной машине



Выбор ОС









- Лучше ставить на Linux
- Рекомендуется CentOS (7 build 1503)
- Под Windows ставить чуть сложнее (надо выкачивать дистрибутивы)
- Под Windows могут быть проблемы с PostgreSQL
- USMF – под Windows Server 2008 R2



Установка ОС и утилит

- CentOS минимальная (не сетевая!)
- Утилиты: `ssh2` сервер + `nano` + `mc` + `gpm`
- Web сервер (Apache, ...) не нужен, если не хотите использовать PgHPAdmin
- Графическая оболочка не нужна

Установка поддержки

- Java (Oracle/Open JDK) ≥ 7 (8 update 66) 
- Tomcat ≥ 7 (8.0.28) (и др.) 
- PostgreSQL ≥ 8.4 (9.4.5) 
или Oracle $\geq 10g$ 
- maven 
- ant 
- maven ≥ 3 + ant ≥ 1.8 – есть в CentOS 7



Установка DSpace

- Не сложная, UNIX-стиль. Важно правильно создать пользователей, группы, раздать права, запускать от нужного пользователя.
- mvn package – подгружаются модули на Яве и формируется задание на сборку, система может работать
- Останавливается Tomcat (система недоступна), и система собирается (ant с параметрами fresh_install, update, etc.)
- Запустить кота
- В первый раз создать суперадмина
- Иногда можно (Windows) править на живую, даже без остановки кота

- Регистрация хотя бы на сайте dspace.org (24.10.2015 – 2844, 123 страны, Молдова – 1). coar-repositories.org – Confederation of Open Access Repositories
- Бэкапы (локальные изменения, assetstore, БД, система)
- Реиндексация для поиска и просмотра
- Автоматизация (напр., ежедневная реиндексация)
- Развитие – возможны локальные адаптации, переводы, дополнительные метаданные
- Обновление – на версию 5 можно перейти с любой меньшей. Ставится система и реструктурируется БД.



Выводы

- Установка простая
- Эксплуатация и поддержка – простые
- Требуются периодические бэкапы
- Требуется периодическая индексация

Спасибо за внимание!



Institutional Repository in Medical Sciences
(IRMS – Nicolae Testemițanu SUMPh)



[University homepage](#) | [Library homepage](#)

Search DSpace

Go

[Advanced Search](#)

[Home](#)

Browse

- [Communities & Collections](#)
- [Issue Date](#)
- [Author](#)
- [Title](#)
- [Subject](#)

Sign on to:

- [Receive email updates](#)
- [My DSpace authorized users](#)
- [Edit Profile](#)
- [Help](#)
- [About DSpace](#)

IRMS - Nicolae Testemițanu SUMPh >



Welcome to the Institutional Repository in Medical Sciences of the Nicolae Testemițanu State University of Medicine and Pharmacy of the Republic of Moldova.

IRMS – SUMPh collects, preserves and gives open access to conference papers, articles from scientific journals, reports, theses and dissertations, study aids, patents, student theses and other publications created by members of the Nicolae Testemițanu University of Medicine and Pharmacy.

Search

Enter some text in the box below to search DSpace.

Go

Communities in DSpace

Choose a community to browse its collections.

- [BIBLIOTECA ȘTIINȚIFICĂ MEDICALĂ a IP USMF "Nicolae Testemițanu"](#)
- [COLECȚIA INSTITUTIONALĂ](#)
- [FACULTATEA FARMACIE](#)
- [FACULTATEA MEDICINĂ](#)
- [FACULTATEA REZIDENTIAT ȘI SECUNDARIAT CLINIC](#)
- [FACULTATEA STOMATOLOGIE](#)
- [REVISTE ȘTIINȚIFICE ALE IP USMF "NICOLAE TESTEMIȚANU"](#)
- [ȘCOALA DOCTORALĂ](#)



RSS Feeds

[RSS 1.0](#) [RSS 2.0](#)

[RSS](#)



Tempus



[DSpace Software](#) Copyright © 2002-2010 [Duraspace](#) - [Feedback](#)