

Е-архиви: астрономически фотографски плаки и български народни песни

Николай Киров

Департамент Информатика, НБУ
Институт по математика и информатика, БАН

Е-АРХИВИ ЗА Е-НАУКА

24.10.2011

В помощ на научноизследователския процес:
Манифест на академичните библиотеки

Библиосфера 12 (2010)

- Крис Бърг, Станфордски университет
- Рос Колеман, университет Сидни
- Рики Еръуей, изследователски отдел, OCLC (Online Computer Library Center)

Призив за действие

За да продължат да играят важна роля за научноизследователския процес, **АКАДЕМИЧНИТЕ БИБЛИОТЕКИ ТРЯБВА ДА:**

9. Заинтригуват учените с представяне на основни научни данни, които се нуждаят от дългосрочно съхраняване и организиране на достъпа.

Е-архиви: астрономически фотографски плаки и български народни песни

Съдържание

- Е-наука и ролята на академичните библиотеки
- Е-архив: астрономически фотографски плаки
- Е-архив: български народни песни

За ролята на академичните библиотеки

За съдържанието на понятието *библиотека*

- Класическа библиотека → Дигитална библиотека
- Класически архив → Е-архив
- Класическа наука → Е-наука
- Класическо обучение → Е-обучение

За ролята на академичните библиотеки

Класическа библиотека – използване

- Описание (метаданни) = Каталози
- Търсене = Хартиени каталози + Е-каталози
- Данни = Книги
- Намиране + достъп = На рафта според класификацията
- Използване = Заемна служба, четене

Дигитална библиотека – съхранение и използване

- Описание (метаданни)
- Търсене = Е-каталози + Google
- Данни = Е-книги (текст) + образ + музика + данни
- Достъп
- Използване = Четене + гледане + слушане + *компютърна обработка*

За ролята на академичните библиотеки

Класически архив – съхранение

- Описание (метаданни)
- Ограничен достъп (за наука, за политика, ...)

Е-архив – съхранение и използване

- Описание (метаданни) – задължително
- Публичен достъп

За ролята на академичните библиотеки

Класическа наука

- *“На раменете на гигантите.”*
- Експерименти (реални) – данни
- Публикации

Е-наука

- Какво е направено до сега по дадена тема?
- *Данни*
- Публикации и Е-публикации

Примери

- Астрономия
- Фолклор
- Български език

Е-архив: астрономически фотографски плаки

Проект *Астроинформатика: Обработка и анализ на дигитализирани астрономически данни и web-базирано приложение,*

финансиран от ФНИ на МОМН, 2009-2012. В проекта участват специалисти от 3 института на БАН, от два университета (СУ и НБУ), от Военна академия, от Чехия и Германия.

- Астрономическите фотографски плаки са снимани с различни камери, монтирани на телескопи от 70-те години на 19 век до 90-те години на 20 век.
- Общо по света има малко над 2 милиона плаки, получени на около 400 професионални телескопи.
- Почти всички астрономически открития до сега са направени по фотографски плаки.
- Астрономите измерват местоположението и яркостта на обекта (звезда, галактика, планета) на плаката.

Днес старите (дигитализирани) плаки се използват:

- за определяне на точните орбити на астероиди и комети,
- за изследвания на променливи обекти,
- за моделиране на дълговременни промени на озоновия слой,
- за изследване на движенията на земната ос,
- ...

База данни от широкоъгълни астрономически плаки (WFPDB)
Създадена е информационна технология за описание на плаки (над 50 параметри), в момента 370 хил. плаки са описани в База данни от широкоъгълни астрономически плаки (на свободен достъп, wfpdb.org), единствена в света.

Проблеми

- Класификация
- Избор на метаданни (WFPDB, FITS header)
- Кой да го направи?
- Изгубена информация (Logbook)
- *Астроархеология*

EUROPEAN plate archives digitization last 5-7 years

Place	TNP	Scanner	SNP	Res.
Konkoly	13000	UMAX PL3000	500/500	8/20
Brussels	20000	ROB Agfa DUOSCAN	≈600	250
Sofia	10000	EPSON1640XL/V700	1035/500	16/10
Moscow	20000	GAISH CREO	≈400	10
Moscow INASAN	4000	2xEPSON1640XL	200	16
Bucharest	12000	Umax AlfaVista II	100	10
Cluj	5000	HP	200	250
Vatican	10000	EPSON1640XL	3000	16
Belgrade	12000	EPSON V700	3000	test

TNT – Total Number of Plates

SNP – Scanned Number of Plates

Res. – Resolution in microns/pixel

Totally: more than 250 000 scanned for last 5-7 years.

The new US plate digitization projects

- Harvard College Observatory: 500 000 plates
 - 2004 NSF granted 0.5 million USD for special scanner construction (in 2007 ready).
 - For plate digitization Harvard asked more 5 million USD for the next 5 years.
 - Estimated amount of plate digitization – 75 000-150 000 per year.
 - Now they have 18 000 scanned plates in 5 selected areas.

Технология за дигитализация на плаки

Сканиране и обработка на получените файлове:

- preview (color image, jpg)
- scan (16 bit grayscale, tif)
- FITS header (text)
- grayscale step wedge
- tif2fits

Организация за съхранение и достъп до големи файлове.

Е-архив: астрономически фотографски плаки

Пример.

За плака с WFPDB идентификатор K0N060 005018, системата за дигитализация създава няколко файла с префикс K0N060_005018 и следните суфикси:

Extension	Size	Type	Produced by	Comment
.jpg	1.4M	color image	–	preview
.tif	477M	16bit gs image	VueScan	scan
.hdr	6K	text	header2011	header
.fits	440M	16bit gs image	tif2fits	plate
.hdrf	7K	text	tif2fits	header
w.fits	38M	16bit gs image	tif2fits	wedge

Logbook

Когато има дневник за наблюденията, той се сканира и записът в този дневник за дадена плака може да се види от връзка с идентификатора на плаката.

Проект *Информационни технологии за представяне с музика, ноти и текст на български народни песни в дигитална библиотека,*

частично финансиран от ФНИ на МОМН. В проекта участват специалисти (информатици, музиколози, фолклористи, българисти, музиканти) от:

- Институт по математика и информатика, БАН
- Институт за етнология и фолклористика с Етнографски музей, БАН
- Институт за изследване на изкуствата, БАН
- Факултет по математика и информатика, СУ
- Нов български университет

nikolay.kirov.be/2010/folk/FOLK.html

www.math.bas.bg/~nkirov/2010/folk/FOLK.html

Източник: архив на проф. Тодор Джиджев

Съдържа над 1000 песни, записани през 60-те и 70-те години на XX век.

- над 2000 страници с написани на ръка ноти и на пишеща машина текстовете на песните
- 27 стари магнетофонни ленти с оригинални изпълнения на песните
- 13 тетрадки и тефтери с теренни записи

Дигитална библиотека

Основен обект – песен (идентификатор) – елементи:

- ноти – LilyPond кодиране (1100)
- музика – MP3 файлове, след професионална обработка (1300)
- текст – L^AT_EX формат, с метаданни (1100)
- ръкопис – сканирани в JPG формат (540)

Динамични свойства:

- Добавяне на основни обекти
- Добавяне на нови метаданни за съществуващи обекти
- Добавяне на средства за работа с основни обекти

Е-архив: български народни песни

Използване на дигиталната библиотека

За Е-наука:

Музиколози, българисти, фолклористи, ...

- търсене в текст и ноти
- индекси
- подобие

Пълнотекстова търсачка (Demo)

“За всеки”:

Изпълнители на народни песни, композитори, образование

- печат
- мултимедия

Мултимедийно представяне (Demo)

Е-архиви: астрономически фотографски плаки и български народни песни

Благодаря за вниманието!

- nkirov@nbu.bg
- www.wfpdb.org
- nikolay.kirov.be/2010/folk/FOLK.html

-
-
-
-