

## КЪМ РЕГИОНАЛЕН АГРЕГАТОР НА ЦИФРОВИЗИРАНИ КУЛТУРНО-ИСТОРИЧЕСКИ ОБЕКТИ

**доц. Елена Сомова<sup>1</sup>, н.с. Георги Врагов<sup>2</sup>,  
проф. Георги Тотков<sup>1</sup>**

<sup>1</sup>ПУ „Паисий Хилендарски“, ул. „Цар Асен“ 24, гр. Пловдив 4000

eledel@uni-plovdiv.bg, totkov@uni-plovdiv.bg

<sup>2</sup>БАН, Институт по математика и информатика, гр. Пловдив 4000, бул. „Руски“ 139  
vragov@gmail.com

**Резюме:** Разгледани са въпроси, свързани с дигитализация, създаване, съхранение и разпространение на цифрови културно-исторически обекти. Представени са съвременни стандарти, използвани в областта. Във връзка с инициатива Еурогеапа е проучено състоянието на процеса на цифровизация на регионалното културно-историческо наследство. Коментира се възможността за създаване на регионален агрегатор на цифрови културно-исторически артефакти, базиран на стандарта, използван от Еурогеапа.

**Ключови думи:** цифровизация на културно-исторически обекти, Еурогеапа, стандарти за културно-исторически обекти

### 1. Въведение

Информационните технологии се прилагат широко при събиране (вкл. възстановяване), съхраняване, обработване и разпространение на информация за културно-исторически обекти (КИО). Информацията за културно-историческите обекти се съхранява в цифрови документи под формата на данни, разположени върху различни медии (с локален или глобален достъп), организирани в статични Интернет-страници, бази данни или цифрови хранилища.

**Събирането на информация (данни)** за културно-историческия обект включва неговата дигитализация и въвеждане на съответните метаданни. Дигитализацията представя културно-историческия артефакт в дигитален формат, а метаданните (структурирана информация за дигитализирания обект) се използват за различни операции върху него - идентификация, описание, управление, осигуряване на достъп, и т.н.

При **обработване на КИО**, особен интерес представляват извличането на съответни метаданни и т.нар. *оперативна съвместимост* между различни описания (системи метаданни) за КИО. Оперативната съвместимост между системи от метаданни дава възможност - процедури и обекти, свързани с

обработване на дигитални КИО, да се управляват по един и същи начин, като се предоставят възможности за обмен на данни между различни системи. При проектиране и създаване на софтуерни системи за обработка на дигитални КИО е необходимо да се осигури не само *техническа съвместимост* (свързана с ползване на общи технически стандарти - файлови типове, метаданни и др.), но и *семантична съвместимост* – с ползване на общи тезауруси (речници за използваната терминология).

**Разпространението** на дигитални КИО може да се реализира чрез различни средства: презентационни сайтове; виртуални каталози; виртуални турове; виртуални музеи, библиотеки и архиви; информационни културно-исторически портали; бизнес-приложения (в областта на туризма, аукционите, колекционирането, криминалистиката и др.), и т.н.

## 2. Стандарти за цифровизирани КИО

При описание на КИО се използват множество стандарти. Повечето от тях предоставят XML-схема с основни метаданни, които описват цифровизираните обекти. Според областта на приложение се разработват: общи и специфични стандарти за КИО; подходящи за откриване, търсене и извличане на цифрови ресурси; предназначени за различни области – музеи, библиотеки, архиви, и др.

**Стандартите за музеи** предоставят подходящи системи от метаданни за описание на музейни обекти. Стандартът Categories for the Description of Works of Art - CDWA [3], създаден от College Art Association (САЩ, 1990 г.) се състои от 31 категории с общо 505 метаданни за описание на произведения на изкуството (обекти и изображения). Стандартът има и олекотена версия CDWA Lite. Стандартът FDA/ADAG Guide to the Description of Architectural Drawings към CDWA включва допълнителни метаданни за архитектурни обекти.

**Стандарт Museumdat** [25] е създаден от Institut für Museumsforschung SMB-PK (Zuse-Institut Berlin през 2006 г.) за извличане и автоматично публикуване на основни метаданни в музейни портали. Стандартът е обобщение на CDWA Lite и се състои от 5 категории със 114 метаданни.

Стандартът **Object ID** [28] на J. Paul Getty Trust от 1999 г. представлява малко подмножество на CDWA.

Стандартът **SPECTRUM** [29] е разработен от музеи във Великобритания през 2007 г. Поради това, че стандартът е доста обемист (съдържа 481 метаданни), за малките музеи е разработен олекотен вариант SPECTRUM Essentials. Освен метаданни, SPECTRUM съдържа и описание на 21 музейни процедури, придружен от необходимите поддържащи данни.

Light Information Describing Objects (**LIDO**) [17] е нов стандарт от 2009 г., създаден на базата на CDWA Lite, CIDOC CRM, museumdat и SPECTRUM, и се

състои от 12 категории със 75 метаданни. Стандартът е използван от проект Athena.

**FDAGuide** [12] на Foundation for Documents of Architecture от 1994 г. е предназначен за описание на архитектурни документи и съдържа 92 метаданни, разпределени в 5 категории.

**Стандартите за архиви** предоставят общи метаданни за архивни описания - независимо от физическата медия, върху която са разположени документите.

General International Standard Archival Description - **ISAD(G)** [15] е стандарт от 1994 г. на International Council on Archives (Канада), и съдържа 26 метаданни, разпределени в 7 категории.

International Standard Archival Authority Record for Corporate Bodies, Persons and Families [14] е аналогичният стандарт на ISAD(G) за Австралия, разработен от Committee on Descriptive Standards през 2003 г.

Describing Archives: A Content Standard (**DACS**) [5], приет от Society of American Archivists през 2004 г., е американският аналог на ISAD(G) и ISAAR(CPF). DACS съдържа 31 метаданни в 10 категории, като представя множество от правила за описание на архиви, лични документи и колекции от ръкописи.

Encoded Archival Description - **EAD** [7] на Society of American Archivists и MARC Standards Office of the Library of Congress е стандарт от 2002 г. за описание на архиви и колекции, и за кодиране на документи.

Reference Model for an Open Archival Information System - **OAIS** [27] е стандарт на Management Council of the Consultative Committee for Space Data Systems от 2002 г., който обобщава Z39.50 Profile.

**Стандартите за библиотеки** дават възможност за поддържане на стандартизирани библиографски описания.

Стандартът Functional Requirements for Bibliographic Records - **FRBR** [13] на International Federation of Library Associations and Institutions от 1998 г. представлява 'обектно-свързан модел' от метаданни за библиографски описания.

**MAB2** [18] е стандартът на German National Library от 2001 г. и съдържа много подстандарт като стандарт за: библиографски данни, лични имена, корпоративни имена, заглавия, локални данни, адреси и библиотечни данни и данни за класификация и нотация.

MARC 21 (**MARC**) [19] е може би най-известният и разпространен стандарт, състоящ се от множество подстандарт, разработени от Library of Congress от 1999 г. MARC съдържа 5 подстандарт за Bibliographic Data, Authority Data, Holdings Data, Classification Data и Community Information.

Стандартът Metadata Encoding & Transmission Standard - **METS** [20] е създаден от Digital Library Federation през 2007 г. с 33 XML елемента, разположени в дървовидна структура със 158 атрибута. Поддържа се в Network Development и MARC Standards Office на Library of Congress, САЩ.

**MODS** [24] е стандарт на Library of Congress от 2008 г., като предлаганата XML-схема може да се използва и за други цели освен за библиографски описания.

**MIDAS** [23] може да се определи като **специфичен стандарт за описание на историческото наследство**. MIDAS е разработен от English Heritage през 2008 г. за документиране на сгради, археологически обекти, корабкрушения, артефакти и др.

**Общите стандарти** могат да се използват в различни области. CIDOC Conceptual Reference Model (**CIDOC-CRM**) [4] е създаден през 2006 г. като ISO стандарт за документиране в архиви, библиотеки и музеи. Стандартът предоставя обектно-ориентиран модел със 148 йерархични класове.

**VRA Core** [32] е стандарт на Visual Resources Association's Data Standards Committee от 1982 г. за описание на визуални обекти от културното наследство, тясно свързан с CCO и CDWA.

Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images (**CCO**) [2] на Visual Resources Association от 2006 г. е на основата на VRA Core 3.0 и CDWA, като съдържа 13 категории метаданни със 119 елемента. Използва се за музеи, колекции с визуални ресурси, архиви и библиотеки за изкуство и архитектура, археологически места и др.

Друга група стандарти са **стандартите за откриване на ресурси**. Dublin Core (**DC**) [6] е безспорно най-разпространеният стандарт, разработен от Dublin Core Metadata Initiative през 1995 г. Стандартът съдържа в основната си част (dc namespace) само 15 елемента: contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title и type. Освен това притежава и допълнителни 7 групи със 126 метаданни.

Последната група са **стандартите за търсене и извличане на ресурси**. Стандартът **OAI-PMH** [26] на Open Archives Initiative от 2002 г. представлява протокол за събиране на метаданни и е пряко свързан с DC и XML. Езикът за заявки към релационни бази данни **SQL** е ISO стандарт от 1987 г. **Z39.50** [33] на Consortium for the Computer Interchange of Museum Information от 1988 г. е протокол за търсене и извличане, който съдържа 9 обекта, като 69 тага само в единия обект CIMI Tag Set.

### 3. Цифрови хранилища

С проекта **TEL** (The European Library: Gateway to Europe's Knowledge) [30] от 2001-2004 г., стартира инициатива за създаване на Европейска Цифрова Библиотека (**EDL**) [8]. През 2005 г. започва да действа портал на виртуална

библиотека, която днес предлага достъп до ресурсите на 47 европейски национални библиотеки на 35 езика. EDL предлага търсене и извличане на метаданни и дигитални обекти (свободно или срещу заплащане) на книги, списания, вестници, аудио записи и други материали. TEL използва стандартът DC с някои разширения и е съвместим с Z39.50, MARC 21, UNIMARC и ISO 2709.

Следващи проекти разширяват EDL: **TEL-ME-MOR** (The European Library: Modular Extensions for Mediating Online Resources) в периода 2005–2007 г.; **EDLproject** [9] в периода 2006–2008 г.; **TELplus** [31] в периода 2007–2009 г. и **FUMAGABA** в периода 2008–2009 г. Проектът TELplus е и с българско участие (Народна библиотека “Св. св. Кирил и Методий”).

През 2007 г. стартира инициативата **Europeana** [10] (последовател на TEL) с основна цел – създаване на (многоезичен) достъп до културното-историческото и научно наследство на Европа, вкл. цифрово съдържание на архиви, библиотеки, музеи и аудио-визуални колекции. За описание на обектите се използва стандартът DC, допълнен с няколко специфични метаданни на Europeana – общо 49 метаданни. От тях 7 са силно препоръчителни, 10 - препоръчителни, 20 - допълнителни и 12 - специфични.

Уебсайтът на Europeana стартира през ноември 2008 г., като до края на 2009 г. вече съдържа 4.6 милиона обекта. До юни 2010 г., броят цифровизирани обекти на Europeana трябва да достигне 10 милиона, а до 2015 г. - 15 милиона. От 2011 г. уебсайтът ще има и семантични компоненти.

В момента Europeana има 108 партньори от 23 страни, като продължава допълването ѝ с помощта на нови проекти, свързани със създаване на регионални и локални **агрегатори на цифрови КИО**. Така напр., проекти **Multilingual Inventory of Cultural Heritage in Europe - MICHAEL** [21] (в периода 2004–2008 г.) и **MICHAELplus** [22] (2006–2009 г.) са свързани с агрегатори на Europeana, които предоставят многоезично описание на цифрови ресурси. За записване на съответните метаданни се използва DC с някои разширения – MICHAEL-EU Dublin Core Application Profile, който съдържа 147 метаданни. В MICHAEL участват 16 български институции, вкл. Регионалният исторически музей и Археологическият музей, гр. Пловдив.

Проектът **EuropeanaLocal** [11] (2008 г.-2011 г.) подпомага включване на локални и регионални библиотеки, музеи, архиви и аудио-визуални архиви към Europeana. Проектът има голяма партньорска мрежа от регионални и местни институции на 27 държави, а за описание на обекти се използват стандартите на Europeana.

Проектът **Judaica Europeana** [16] има за цел да осигури достъпа до европейската еврейска култура.

Друг агрегатор на Europeana е проектът **Access to Cultural Heritage Networks Across Europe (Athena)** [1] (2008 г.-2011 г.). За описание на обектите се

използва стандарт LIDO. В проекта участват 24 държави (общо 120 институции). България участва с 5 колекции на Централната библиотека към БАН.

#### **4. Към регионален агрегатор на културно-историческото наследство**

Еуропеана се превръща в най-мощното европейско цифрово хранилище на обекти на културно-историческото наследство. Естествено е усилията на съответните институции в гр. Пловдив да се насочат към присъединяване към Еуропеана. Включването (според приетата практика) става на два етапа: обединяване на ресурси на няколко институции и оформяне на т. нар. агрегатор, и след това - предоставяне на агрегираното цифрово съдържание на Еуропеана. За описание на КИО, Еуропеана използва малко на брой метаданни, поради което е добра основа за създаване на цифрови хранилища за различни по тип и описания обекти на музеи, галерии, архиви и библиотеки.

В изследването е проучено състоянието на процеса на дигитализация и съхранение на цифрови КИО в гр. Пловдив с цел – установяване на степента на готовност за регионално агрегиране към Еуропеана.

Библиотеките в гр. Пловдив, както и в страната, използват MARC-стандарта и имат разработени виртуални каталози за библиографските описания. В регионалната Народна библиотека „Иван Вазов“ са дигитализирани няколко книжни колекции. По проект на МОН и Световната банка, от 2002 г., шестте пловдивски библиотеки на държавните университети имат общ виртуален своден каталог

Процесите на дигитализация и описания на КИО в пловдивските музеи са на различен стадий от своето развитие. Съответната информация, обаче, в по-голямата си част се разполага в бази данни и електронни каталози, които не следват европейски стандарти (пример – Регионалният исторически музей).

Регионалният археологически музей използва Excel-ски таблици за представяне на археологическите обекти, следвайки MARC-стандарт. Библиографските описания на библиотеката на музея са вече в база от данни, а до края на годината се очаква и табличните данни да бъдат прехвърлени.

Пловдивските галерии не разполагат с цифрови варианти на притежаваните и представяните колекции от картини.

Етнографският музей в гр. Пловдив поддържа 2 (два) вида документи в електронен формат: паспорти на обектите в Word-формат и инвентарни книги в Excel-формат, от които автоматизирано или ръчно биха могли да се извлекат стойности за метаданните на Еуропеана. В Табл. 1. е представен анализ – дали (и кои) данни за КИО на Етнографския музей (гр. Пловдив) могат да бъдат използвани за създаване на метаданни за Еуропеана. Със знак 'И' е указано, че съответните метаданни са на разположение (в един от двата документа), с 'Н' – метаданните не са налични, а с 'А' – метаданни, които биха

могли автоматично да се създадат при съхраняване на КИО в съответното цифрово хранилище.

При прехвърляне на данни за КИО на Етнографския музей, например, в хранилище, отговарящо на стандартите на Europeana, ще се наложи допълнително 'ръчно' въвеждане (например в случая, когато от едно поле на музеен документ трябва да се 'извлекат' две метаданни на Europeana).

**Таблица 1. Съответствие между метаданни (Europeana) и данни за КИО на Етнографския музей (гр. Пловдив)**

| Силно препоръчителни  |   | Допълнителни            |   | Europeana           |   |
|-----------------------|---|-------------------------|---|---------------------|---|
| dc:title              | И | dc:format               | А | europeana:country   | А |
| dcterms:alternative   | И | dcterms:extent          | И | europeana:hasObject | А |
| dc:creator            | И | dcterms:medium          | И | europeana:isShownAt | А |
| dc:contributor        | И | dc:identifier           | И | europeana:isShownBy | А |
| dc:date               | И | dc:rights               | А | europeana:language  | А |
| dcterms:created       | И | dcterms:provenance      | И | europeana:object    | А |
| dcterms:issued        | А | dc:relation             | И | europeana:provider  | А |
| <b>Препоръчителни</b> |   | dcterms:conformsTo      | А | europeana:type      | А |
| dc:coverage           | И | dcterms:hasFormat       | А | europeana:unstored  | И |
| dcterms:spatial       | И | dcterms:isFormatOf      | Н | europeana:uri       | А |
| dcterms:temporal      | Н | dcterms:hasVersion      | Н | europeana:usertag   | А |
| dc:description        | И | dcterms:isVersionOf     | Н | europeana:year      | А |
| dcterms:isPartOf      | И | dcterms:hasPart         | И |                     |   |
| dc:language           | А | dcterms:isReferencedBy  | Н |                     |   |
| dc:publisher          | И | dcterms:references      | Н |                     |   |
| dc:source             | И | dcterms:isReplacedBy    | Н |                     |   |
| dc:subject            | И | dcterms:replaces        | Н |                     |   |
| dc:type               | И | dcterms:isRequiredBy    | Н |                     |   |
|                       |   | dcterms:requires        | Н |                     |   |
|                       |   | dcterms:tableOfContents | И |                     |   |

Предстояща задача е създаването на регионален агрегатор на цифрови КИО, базиран на стандарта, използван от Europeana. Предполага се агрегаторът да е достъпен за всяка регионална културно-историческа институция (за съхраняване на дигитализирани КИО) и за посетители (за търсене, разглеждане и справки). Създаването на регионален агрегатор е първа стъпка към представяне и популяризиране на богатото културно-

историческо наследство на гр. Пловдив и региона в Европейското цифрово пространство.

Разработката е частично финансирана по проект Д002-308 „Автоматизирано генериране на метаданни за спецификации и стандарти на е-документи“ към Националния фонд за научни изследвания.

### Литература

- [1] Athena, <http://www.athenaeurope.org/>.
- [2] CCO, <http://www.vrafoundation.org/ccoweb/cco/intro.html>.
- [3] CDWA, [http://www.gettytrust.us/research/conducting\\_research/standards/cdwa/](http://www.gettytrust.us/research/conducting_research/standards/cdwa/).
- [4] CIDOC-CRM, <http://cidoc.ics.forth.gr/>.
- [5] DACS, <http://www.archivists.org/governance/standards/dacs.asp>.
- [6] Dublin Core, <http://dublincore.org/>.
- [7] EAD, <http://www.archivists.org/saagroups/ead/aboutEAD.html>.
- [8] EDL, <http://www.theeuropeanlibrary.org/portal/organisation/cooperation/> archive/edlproject.
- [9] EDLProject, <http://www.edlproject.eu>.
- [10] Europeana, <http://www.europeana.eu/>.
- [11] EuropeanaLocal, <http://www.europeanalocal.eu/>.
- [12] FDA Guide, [http://www.getty.edu/research/conducting\\_research/standards/fda/](http://www.getty.edu/research/conducting_research/standards/fda/).
- [13] FRBR, <http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records>.
- [14] ISAAR (CPF), [http://www.icacds.org.uk/eng/ISAAR\(CPF\)2ed.pdf](http://www.icacds.org.uk/eng/ISAAR(CPF)2ed.pdf).
- [15] ISAD(G), <http://www.ica.org/en/node/30000>.
- [16] Judaica Europeana, <http://www.judaica-europeana.eu/>.
- [17] LIDO, [www.athenaeurope.org/getFile.php?id=535](http://www.athenaeurope.org/getFile.php?id=535).
- [18] MAB2, <http://www.d-nb.de/eng/standardisierung/formate/mab.htm>.
- [19] MARC, <http://www.loc.gov/marc/>.
- [20] METS, <http://www.loc.gov/standards/mets/>.
- [21] MICHAEL, <http://www.michael-culture.org/bg/home>.
- [22] MICHAELplus, <http://www.michael-culture.org/bg/home>.
- [23] MIDAS, <http://www.english-heritage.org.uk/server/show/nav.8331>.
- [24] MODS, <http://www.loc.gov/standards/mods/>.
- [25] Museumdat, <http://museum.zib.de/museumdat/museumdat-v1.0-en.pdf>.
- [26] OAI-PMH, <http://www.openarchives.org/>.
- [27] OAIS, <http://public.ccsds.org/publications/archive/650x0b1.pdf>.
- [28] Object ID, <http://icom.museum/objectid/>.
- [29] SPECTRUM, <http://www.collectionstrust.org.uk/spectrum>.
- [30] TEL, <http://www.theeuropeanlibrary.org>.
- [31] TELPlus, <http://www.theeuropeanlibrary.org/telplus/>.
- [32] VRA Core, <http://www.vraweb.org/projects/vracore4/index.html>.
- [33] Z39.50, <http://www.unt.edu/wmoen/projects/CIMI/HarmonProfile1.htm>.