Vedrana Juričić Croatian Academy of Sciences and Arts, Library

DIGITAL COLLECTION OF CROATIAN ACADEMY OF SCIENCES AND ARTS

Abstract: The Digital collection of the Croatian Academy of Sciences and Arts (http://dizbi.hazu.hr) includes digitized materials of the Academy's research units and the Academy Library. The Collection currently comprises the following materials in original formats: books, journals, cast sheets, microfilms, plaster casts, manuscripts of poems and music, paintings, art exhibition catalogs etc. The article discusses the structure, virtual exhibition, viewing, tags and references, searching and other features developed in the repository.

Keywords: digital collection, the Croatian Academy of Sciences and Arts, dizbi, virtual exhibition, Silverlight, JavaScript browser, references, video

1. Introduction

The Digital collection of the Croatian Academy of Sciences and Arts (http://dizbi.hazu.hr) includes digitized materials of the Academy's research units and the Academy Library. It was launched in 2009 with the aim of serving as a unique digital repository, representing the Croatian Academy's scientific and artistic collection on one Web site, offering free Web access and modern and standardized development of technical support.

The following Academy units: The Fine Arts Archives, the Institute of the History of Croatian Literature, Theatre and Music: The Division of the History of the Croatian Theatre and the Division of the History of Croatian Music, and the Academy Library as initiator, have established the Digital Collection. The Glyptotheque joined in the repository in March 2011. Its task is to collect digitized reproductions of major works of statuary art by means of plaster casts. The latest units are the Division of the History of Croatian Literature, The Ethnology Division, The Strossmayer Gallery of Old Masters and The Croatian Museum of Architecture.

The Library director coordinates the team consisting of one or more representatives (mostly younger researches) from each Academy unit. The project is funded from multiple sources as follows: the annual budget of the Academy (140,000 HRK in 2011) for the project of digitization), The Foundation of the Academy (15,000 HRK in 2011), annual competitions for financial support for digitization by the Ministry of Culture of the Republic of Croatia (60,000 HRK in 2011) and of the City of Zagreb Office for Education, Culture and Sports (67,000 HRK). The total amount was 267,000 HRK in 2011, or since the beginning of the project of digitization in 2008 approximately 550,000 HRK.

The Digital Collection currently comprises the following materials in original formats: books and journals (296,054 pages or 889 volumes), art exhibition catalogs (5,504 pages or 258 volumes), cast sheets (7,063), microfilms (1,266 frames or 72 units), plaster casts (237), medals and plaques (171), copies of frescos (10) and Croatian sculptures (238), manuscripts (12,312), photos (74), paintings (100). The total amount is 323,029 digitized units at the end of 2011. The plan is to store video recordings of lectures to be held in the Academy.

The criteria for the selection of materials to be digitized are based upon the principles of scientific, cultural and historical value and interest, preventive protection of old and rare items, and the promotion of the scientific, research, artistic and publishing activity of the Croatian Academy of Sciences and Arts.

The digitized documents are stored in several formats: masters in high resolution 24 bits uncompressed TIFF format (400 dpi), derived images in JPEG format, textual material in PDF additionally processed for online access, following the procedures and regulations recommended by the Croatian Cultural Heritage Project Office. Uploaded documents in different formats like born-digital doc, xls etc. except audio and video documents, are converted on-the-fly to PDF upon upload. The original files and created PDF files are stored on the server as well, but they are not accessible to a public user. The ArhivX concept applied in the repository allows the access to single page image without the need to download the full document on the client side. A user can download document in PDF format with embedded watermark generated on-the-fly.

Conceptual design and programming of the Digital Collections of the Croatian Academy of Sciences and Arts was conducted by the company ArhivPRO d.o.o. I thank Mr. Kristijan Crnković for his suggestions and good cooperation.

2. Structure of the repository

In the digital collection of the Academy the arrangement of material, design, viewing and searching should be simple and clear. These are the main objectives pursued.

On the home page of the repository there is the main menu with a list of Academy research and museum-gallery units and a list of the most popular, most read and recently added documents.

3. Virtual exhibition

Virtual exhibition is the only, but striking, visual content on the first page. In the short time intervals various images from the virtual exhibitions are displayed on the homepages of each unit of the Academy. At this level it is possible to access the introductory description of the digitized collection of each Academy unit, and the menu with the list of its digitized collections. (Fig. 1)



Fig. 1. Home page

Virtual exhibition can be found on the home page of each Academy unit. The exhibition presents the most attractive digitized fragments. In the same way it is possible to organize thematic exhibitions.

4. How to display the digitized documents?

There are two ways in which the material is displayed on the first level. The most comprehensive publication of journal / series Rad HAZU, which has a total of 520 volumes retained the original disposition of the main numbering string that is the volume number, without displaying the metadata. (Fig. 2)

😸 Wenebika akademiya (nanarili i mujetnesti - digilala	substantian a	- Meyett	a Finales	ē. –										22
Ge En Jes Helpst Stopping Jog Ref														
Truble states parol (method -d _ T														-
Contraction of the second											2 C (2)	a		1.
🚹 Gaugie Presidinty 🎴 Most Valent 🗌 Gating Started 🗽 Late	si Hendleren 1	lister	Sal - Prijeve	a Balle	Japobi 🖣	Gauge (treated	n stadenije	4. 7. 8	realsha ata	denija a	C RATKO	photos	
🔶 Vale petragina stratica							Rad							
	Hajatariji Akadem	i najipi jina publi	indonji Akr	ademijit da. Do de	man jo ciji	i časopish intjene sko	enja s i gra 517	radovinio recorka	iz oñ s	sanatveni	h i unget	ničkih por	evą i epis	
	Taj je bo 201. brij	ų reči od is, ututer	Araja naji Kepeg čali	ionijeg ce tri mesk	eska, jer i a noter isti	la anecieco beoj 274,	20. st.	sbjæljen	a des eize	Radava	dettične	кленкаў	e od 270. do	
	+ 00	logenhale	npie iper	nika.										1
	1 (1997)	2	3 (1880)	4	8 (1963)	8 (1968)	7 0880	# (1998)	a own	10 () (11)()	in date	10 [1570]	48 (1870)	
		15 (1111)	H	1071)	-	Fail Brig 5 Gedan the	TD	21 (1071)	22 (1107 3)	23 0 8736	2M O OFTED	28 (1073)	28 (1874)	
	27 (1874)	29 0114	29 (1874)	38 (1975)	34 11875	1.4.	in	34 (1076)	96 042%	36 0.678)	NP CATED	38 (1971)	39 11671)	
		48 (1817)	42 (1610)	40 (1920)	44		in	47 (1929)	40 (1828)	49 (1818)	50 0100	51 (1980)	52 (1993)	
	53 (1090)	58 (7 1010)	55 (1880)	54 (11011)	57		deo	48 [1182]	-61 (1882)	62 (1812)	63 (1882)	44 (1982)	85 (1993)	
	886 (1586 (1	87 (1883)	88 0.660	48 (1584)	70 (1884)	78 (1894)	72 0880	73 (1884)	74 (1985)	28 (1888)	78 0.000	177 (1085)	78	
	79 (1963)	80 (1805)	81 (1000)	82 (1085)	43 (1961)	86 (1897)	85 0.007)	88 (1081)	87 (1983)	88 () 800)	85. (1680)	08 (1087)	91 (1900)	
	82 (1993)	82 (1808)	90 (1990)	94 (1988)	95 (1993)	95 (7 808)	97 0.000)	98 (1080)	99 (1880)	108 (7.550)	491 (1.990)	182 (1990)	903	
Mg. Jdds Aanu K. Weler, php?/datest = 1074	104	105	18	187	-	100	191	111	112	113	111	115	115	

Fig. 2. Displaying the most comprehensive publication Rad HAZU of 520 volumes

The initial thumbnail display and associated metadata are used for all other collections on the first level. The arrangement within a single type of document or collection can be easily changed by selecting the key for sorting - author, title, date, etc. (Fig. 3).

Glavni nasl	lov 💌 🕼	4 • 1/4 🕨	
Ì	naslov: Apolon Belvederski vrijeme nastanka: kraj IV.st.pr.n.a broj: inv.br: AP-26	b., rano helenističko razdoblje grčke unjetnosti	1.
	K nastov: Aleksandar Velki autor: Nepoznat vrijeme nastanka: IV. st. pr.n.e., I broj: inv.br. A-104	nelenističko rezdoblje grčke unjetnosti	2.
P 🐴 🎙			

Fig. 3. Displaying by title proper

There are two different shortened viewing interfaces (Fig. 4 and 5). The first emphasizes the metadata and the second image of digitized material, which is suitable for art exhibits - paintings, sculptures, etc.

Glavni	naslov 💽 kr	4 • 1/4	
1.	naslov: Apolon Belvederski autor: Leohar		P ()
2. 🚊	naslov: Aleksandar Veliki autor: Nepoznat		P 0
з. 📱	naslov: Amazonka autor: pripisuje se Polikletu ili Kresilasu		P ()
4.	naslov: Apoksiomen (Atlet sa strugalom) autor: Lizip		P ()
5.	naslov: Ares (ili: Mars) Ludovisi autor: pripisuje se Lizipu ili Skopasu		P ()

Fig. 4. Shortened viewing interface which emphasizes the metadata



Fig. 5. Shortened viewing interface which emphasizes image of digitized material

On the second level the thumbnail image is enlarged. There is also more metadata on this level and (Fig. 6) available tags and references to the related network resources outside of the repository, as well as links to documents that illustrate or complement the main digitized object. For non-linguistic contents there are bookmarks and tags that could be used for indexing purposes as well as for area selection and tagging. For example, some area of a painting can be selected and tagged. (Fig. 7) The system is able to find selected and tagged area through search mechanism.

			Digitalna	zbirka				
Macceud								
	» početna	» o projektu	» pomoć	» KONTAKT	» Hrvatski	» English	» Login	n
						🔎 Pretraživ	/anje 🕕	
	>> De	talji						
DIPLOMATIČKI ZBORNIK semilivne revatske dalavlje i slavonje nazadavlje i slavonje nazadavlje nazadavlje i slavonje nazadavlje nazadavlje i slavenje prepara je na poslovanja za statu slavenje statu slavenje slav	>> De	aslov: 5	tički zbo	ornik Kalje∖	/ine Hr∨a	tske, Da	almaci	ije i Sla∨onije
DIPLOMATIČKI ZBORNIK KALIVIN RVATNEL DALAMIJE I SLAVDNIJE MRA MILAMINE AMAZINE ADALIVIT (UMTYME MILAMINE AMAZINE ADALIVIT (UMTYME) MILAMINE AMAZINE ADALIVIT (UMTYME)	podr jeziki ID: 1	alji Diploma naslov: 5 :: hrv lat 278	tički zbo	ornik Kalje∖	vine Hr∨a	tske, Da	almaci	ije i Sla∨onije
DIPLOMATIČKI ZBORNIK REALDOW RATORE, DALAGJE I SLAVONJE Dala Martines, akademi slavni i domenik Dala se se odrednik odri takov Dala se se odrednik odri takov	P De	talji Diploma naslov: 5 :: hrv lat 278 m objave: 2010-02-	tički zbo	ornik Kalje∖	vine Hr∨a	tske, Da	almaci	ije i Sla∨onije

Fig. 6. The second level of viewing



Fig. 7. In paintings tags lead to its selected parts for iconographic analysis

5. JavaScript browser

On the third level the digitized object can be viewed in JavaScript browser (Fig. 7 and 8). This interface has versatile characteristics. It runs successfully on all platforms such as Internet Explorer, Firefox, Google Chrome, etc., as well as Windows and Linux systems. Accessing the requested page is very fast. A user does not have to wait for the whole multipage document to be downloaded. Instead, the client generates a request for a particular page and sends it to the server. The server responds quickly, while there is a small amount of data that should be transferred to the client side.



Fig. 8. Viewing a document of 784 pages, which are fast accessible through JavaScript browser

The functionality of the JavaScript interface is constantly being improved. The following features exist at the moment: thumbnails and large image preview (the icons 1 and 2), thumbnails only preview viewing (icon 3), large image preview (icon 4), double page preview (icon 5), fit to height (icon 6) fit to width (icon 7), best fit preview (icon 8) and continuously viewing images by vertical sequence (icon 9). Textual materials can be accessed through the table of contents of the journal or book (icon 14). Text extraction from the scanned image is available as well (icon 15). The following are the icons for viewing references (icon 16) and displaying tags (17). The selected range of pages or the entire scanned document can be saved as a .pdf document (icon 18). By clicking on the last two icons you can print the current page (icon 19) or zoom and rotate a picture (icon 20). A

watermark (the official logo of the Academy with the image of its founder, Josip Juraj Strossmayer) appears on every image of an exported or printed document.

6. References

This feature is developed for semantic linking with other repositories that use similar technology and the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) standard.

By clicking on the icon of the JavaScript interface the reference window opens: a click on the right icon displays the section of the related document in our repository. (Fig. 9). By clicking the name of reference a document in another repository that refers to the document in our repository is displayed (Fig. 10).



Fig. 9. Displaying the section of the related document in the Academy repository



Fig. 10. Displaying the document in another repository that refers to the document in the Academy repository

7. Searching

There are integrated searching algorithms for metadata and full-text search supporting the following:

- 1. Full-text or metadata search: use keywords to search the full-text of all items in repository.
- 2. Collection search: search inside a collection of materials
- 3. Single-volume search: when viewing an item in the viewer, use keywords to search inside that volume.

Full-text search supports phrase searching or multiple term searching. Non-linguistic content is searchable through tags and bookmarks set by collection builders.

7.1. Highlighted search result. Search terms are highlighted in the search result list (Fig. 11),

Thumbs Search
Napoleon
1. [str. 381]
da podiže tvornice, da električno svjetlo dovodi u najmanja sela, da na ogromnim brodovima plovi po svi toliko ljudi, toliko glava i toliko kesa sudjeluje u pustolovnim poduzećima? To neograničeno pouzdanje st prije svega u sjaju svoga pobjedonosnoga oružja, [više]
2. [str. 133]
Značajno je, da je medjutim posvema prestao do tada sveobéi upliv francuzke knjige u Holandiji, te narodni patriotski pjesnici Bilderdijk i Tollens. Istina, u srpnju 1810. pripoji <mark>Napoleon</mark> Holandiju Francuzkoj, bratu, koji se već bio "poholandio", no u to [više…]

Fig. 11. Highlighted search result

as well as on the images in the universal viewer (Fig.12).



Fig. 12. Highlighted search result on the image in the universal viewer

As mentioned above searching is supported in a single document, in an entire collection or in all collections of the repository. Searching is available in JavaScript interface or in Silverlight.

Clicking the search icon in the JavaScript viewer of a single document and typing the required term in the search box (Fig. 13) displays the OCR text in a new window with the requested term highlighted. By clicking on the icon at the bottom of the page, the requested page with highlighted search term appears.

	Diplomatički zbornik Kaljevine Hrvatske, Dalmacije i Slavonije		
	· · · · · · · · · · · · · · · · · · ·	# @ 9 8 P	
EGG - Constanting - Constantin	1270, 1. maja. Kaptol stolnobiogradski svjedoči, da je	Kurej sa sinovima svojima	~
P41	proudo palaca a varazaina Dioniz	u javorjuniku kraijue.	
Pretraživanje po dokumentu kurej Korej Asslov: Diplomatički zbornik Kaljevine Hrvatska, D Asslov: Diplomatički zbornik Kaljevine Hrvatska, D Si andržava se	1 almacije i Slavonije, slika 542 1 almacije i Slavonije, slika 542 1 i slavonije, slika 542 1 i slavoni se slavo se	Christi fidelibus presentes ternam. Ad uniuersorum no- quod Cure ciue Worosdiensi una parte, Dionisio officiali dam Thome comitis fratris copi Strigontensis aule regis titutis, idem Curey et fili sui dominus noster Bela illustris nis castri Zaladiensis absque istentem et ad sue potestatis is, uirgultis, fenetis, aquis et cientem eidem Cure eiusque sisus fidelitatis obsequia et	16

Fig. 13. Single volume search

Clicking on the specific collection of documents starts a search within this collection (Fig. 14).

to appears in	Digitalna zbirka		
	» POČETNA » O PROJEKTU » POMOĆ » KONTA	KT » Hrvatski » English » Login	
	Sigismundus	🔎 Pretraživanje 🕕	
OČETNA) Knjižnica) Dij	Pretraživanje po: "Diplomatički zbomik Kraljev Slomatički zbomik Kraljevine Hrvatske, Dalmacije i Slavonije	ine Hrvatske, Dalmacije i Slavonije"	
Pronadeno u textu [239]	Naslov: Diplomatički zbornik Kraljevine Podnaslov: 18 Nađeno na slikama: 545, 536, 540, 547, - 538 215, 1390,12, marti, Sasanundas rov. N dicate invulse datas transumit 320 216, 2398, 28	Hrvatske, Dalmacije i Slavonije, slika 545 538, 552, 525, 526, 528, 537, Icoleo de Gara elusque filis litteras sues de collatios de Gara elusque i a 1397, de eorum statutione in martii[Vi8e]	ne insulae dominium
	2. Noslov: Diplomatički zbornik Kraljevine Podnaslov: 17 Nodeno na slikama: 649, 651, 675, 652, – 641 – 53, 1397, 15, tuli, Zarpisa, Sarri Georgio, filo Ni kola, exertine	Hrvatske, Dalmacije i Slavonije, Slika 649 653, 684, 654, 660, 664, 680, model rez posessionem Pseuch in comitatu Crisi 54, 2387, 21. julii. Zagrabiae. Akus, banuš Sdavoni ecipit, ne a [više]	ensi sitam ae, comiti
	3. Naslov: Diplomatički zbornik Kraljevine Podnaslov: 19 - Dodaci, sv. 1	Hrvatske, Dalmacije i Slavonije, slika 419	Harrison Constant Toronto Constant Marcine Constant Marcine Constant
	Nadeno na slikama: 419, 424, 400, 401, IUPANUS, v. Cemecha 54, v. Crepun 54, v. Dra:	180, 386, 399, 406, il 61 IURAMENTUM Corporale ad sancta Dei evangelij	a 45, 107, harnes

Fig. 14. Collection search

The entire repository search starts by typing a search term on the home page (Fig. 15).

A COLOR	Digitalna zbirka
	» POČETNA » O PROJEKTU » POMOĆ » KONTAKT » Hrvatski » English » Lagin
	kusić 🔎 Pretraživanje 🕕
POČETNA	
Pronađeni autori [1] Pronađeno u textu [32]	L Naslov: Rad, slika 4 Nađeno na slikama: 4, 96, 86, 78, SADRŽAI TABLE OF CONTENTS D Bid. 5, Spaventi. I Padovan. V Čakovac. z. Kurd. N Daković. B. Pokrajac. P Nola. N Ipda. A. Goorginer, Cregouric Netwik. A Soldo-Baid. D. Da Svy, I Despot Melanomas. Interferon Therapy Melanom. Torapia netriferotom 7 D bid. 1 Padovan. N. Ppić. V [vils]
	2. Naslov: Ljetopis, slika 254 Nadeno na slikama: 254, 244, 245, 253, 42, 80, 174, 238, 57, 65, 255 Aspiration Crhology in Children and Adelescents, Coll. Antropol, 34 (1), 2010, 197- 200; 1. Šarnija, J. Liaka, J. Munciferrati, H. Jaulien, J. Alabeg, D. Kovačevć, H. Strum, Z. Kunik, Propositic value of morphthalma-srouded transcription factor and By-resinase as markers for _ (refer)
	2. Nadowi Muzikalije u glazbenom arhivu splitske Stolnice sv. Dujma, slika 381 Nadowi na slikama: 381. - 360 - Sign, LOXXII/1967 ANDHRUS Adiuva nos In a, Crkp part/S, A, T, B/ dini; S, A, T, B/ diplov/Sign. LoXXII/265 ANDHRUS Brannine Pajasmin In S., C rkp dans (T, 1) 11/4 8 1, 3 10 12 v/ 1, 2 v/ 11, 2 v/ 1, 2 v/

Fig. 15. Full-text or metadata search

Advanced search includes the document/collection/repository level and Boolean operators. Search results are grouped by type of metadata (author, title, notes, keywords, etc.), above which is the full text search result as the first listed.

8. Silverlight

Digitized documents sometimes contain high-resolution images which you may want to view in greater detail. This feature is available when using the Microsoft Silverlight plug-in which can be installed by following the instructions on

http://www.microsoft.com/getsilverlight/Get-Started/Install/Default.aspx.

The digital collection of the Croatian Academy of Sciences and Arts includes paintings, sculptures and old and rare books which can be viewed in Silverlight.

Control toolbar includes a toolbar icon for navigation between pages, an icon for increasing or decreasing the size of the viewing window and the pressing the Escape button exits full-screen mode. Zoom in and out of the image using the + / - buttons or the mouse wheel. Rotate the image using the clockwise / counter clockwise arrows. (Fig. 16).

₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽
Sl. 17. G (glagoly)
1. <mark>Kijevski</mark> listići 2. Praški odlomak A 3. Assema- nijev evanđelistar 4. Krčki natpis 5. Bečki listići
6. Bašćanska ploča 7. Senjska ploča 8. Mihano- vićev apostol 9. Humski grafit br. 37 10. Humski
ulomak 13. Legenda o sv. Tekli 14. Ustav 15.

Fig. 16. Interface for viewing the high resolution digitized books with Silverlight

Searching the document displayed in the Silverlight technology is possible as well. In Figure 16. search result is marked in red.



Fig. 17. Video recording

9. Video

Croatian Academy of Sciences and Arts organizes many scientific lectures and conferences, captured regularly in video format and saved on DVDs.

The new video displaying technology can solve the problem of video long term preservation, and at the same time provide access to the recordings to all or selected Internet users. (Fig.17).

10. Using the repository

The Digital collection of the Croatian Academy of Sciences and Arts can be used without authorization. Each user can browse and search without limitation through the entire repository, with the exception of the documents protected by copyright. Such material may only be used in the Academy Library.

Authorized users can create personalized pages, save favorite documents, add tags, etc. The researchers can store their scientific results and working papers during the study on their own personalized pages in the repository.

11. Conclusion

Digital collection of the Academy is being developed very intensively. Since 2008 when the library began developing its digital collection, and since 2009 when the development spread to other Academy units, the repository keeps a total of about 323,000 images of books or separate documents, such as cast sheets, paintings, etc.

Number of Academy units is also growing, from four in the beginning to nine in 2011.

The number of users is also growing very rapidly. In the recent year the repository was visited by about 15,000 visitors, creating bandwidth of 498.0GB.

What's next? We hope that the other Academy units will join into the repository successively and that the process to join Europeana is also an achievable goal in the near future. We are optimistic!

References

- Crnković, K., Juričić, V., Polak Bobić, K. Digitalna zbirka knjižnice Hrvatske akademije znanosti i umjetnosti, Knjižnice: kamo i kako dalje?: zbornik radova 11. dani specijalnih i visokoškolskih knjižnica, Opatija, 1–4. travnja 2009, 2010.
- 2. Vitković, Z. Portal Digitalna zbirka Hrvatske akademije znanosti i umjetnosti, <u>http://www.hkdrustvo.hr/-hkdnovosti/clanak/272</u> [2011-06-20]

- 4. http://daz.hr/bastina/projekti.htm [2011-06-20]
- 5. <u>http://kultura.hr/</u> [2011-06-20]
- 6. <u>http://info.hazu.hr/</u> [2011-06-20]

vedranaj@gmail.com

^{3. &}lt;u>http://dizbi.hazu.hr</u>