

УДК 004.455.032.6XML: 004.738.5

UDC 004.455.032.6XML: 004.738.5

**ПРЕТРАЖИВАЊЕ МУЛТИМЕДИЈАЛНИХ  
ИНФОРМАЦИЈА У СИСТЕМУ  
ДИГИТАЛНИХ БИБЛИОТЕКА  
МАГИСТАРСКИХ И ДОКТОРСКИХ  
РАДОВА<sup>1</sup>****MULTIMEDIA DOCUMENT RETRIEVAL  
IN THE NETWORKED DIGITAL LIBRARY  
OF THESES AND DISSERTATIONS<sup>2</sup>****Бранко Милосављевић, Зора Коњовић  
Факултет техничких наука Нови Сад****Branko Milosavljević, Zora Konjović  
Faculty of Engineering, Novi Sad****Сажетак**

Постоји више софтверских система за претраживање мултимедијалних докумената. Ови системи имају и нека ограничења као што су фиксна структура докумената и ограничен број подржаних типова медија. У циљу превазилажења тих недостатака реализован је Прошириви систем за проналажење мултимедијалних докумената (*Extensible Multimedia Information Retrieval System*, XMIRS). У раду су наведене основне карактеристике архитектуре XMIRS-а која омогућава проширивост система различитим модулима за проналажење мултимедијалних објеката и различитих модела проналажења докумената. Описана је имплементација прототипа XMIRS-а конфигурисаног тако да подржи функције проналажења докумената у софтверском систему Мрежна дигитална библиотека докторских, магистарских и дипломских радова, који се развија на Универзитету у Новом Саду. За имплементацију појединачних захтева искоришћени су постојећи разнородни софтверски пакети који су у XMIRS интегрисани у форми његових модула. На крају су дати илустративни примери претраживања текстуалних садржаја, слика и видео записа.

**Кључне речи: мултимедијални документи,  
проналажење информација, XML документи****1. Увод**

Проблем претраживања мултимедијалних докумената је предмет истраживања у областима проналажења информација (IR) и база података. Појам документа као основног носиоца информација је централни појам у области IR. Третман докумената у разним IR системима је различит, полазећи од неструктурираних текстуалних докумената до структурираних мултимедијалних докумената. Овај рад бави се проблемима претраживања структурираних мултимедијалних докумената. Структурирани мултимедијални документи су сложени објекти који садрже друге објекте. Ин-

**Abstract**

There is a number of systems providing retrieval of multimedia documents. These systems have certain shortcomings, such as fixed document structure and a limited number of supported media types. An *Extensible Multimedia Information Retrieval System* (XMIRS) has been developed as a result of addressing those shortcomings. This paper presents main characteristics of XMIRS software architecture providing system extensibility by multiple multimedia object retrieval modules and multiple document retrieval models. An implementation of a XMIRS prototype configured to support document retrieval in the *Networked Digital Library of Theses and Dissertations* (NDLTD) system is described in detail. Implementation of particular retrieval functionality utilizes existing software packages integrated into the whole system as XMIRS modules. The paper then presents examples of multimedia document retrieval in NDLTD using XMIRS including retrieval of text, images, and video clips.

**Key words: multimedia documents, information  
retrieval, XML documents****1. Introduction**

The problem of retrieval of multimedia documents is a subject of research in information retrieval (IR) and database communities. The notion of a document as a main carrier of information is central in the IR field. The treatment of documents in various IR systems ranges from unstructured textual documents to structured multimedia documents. This paper deals with problems of retrieval of structured multimedia documents. Structured multimedia documents are complex objects comprising other objects. The internal organization of comprised document elements constitutes document structure.

<sup>1</sup> Рад је изложен на *Стручном скупу "Рад у систему узајамне каталогизације"* одржаном у оквиру Девете скупштине Заједнице библиотека универзитета у Србији, 10.- 11. октобра 2003. године у Народној библиотеци Србије.

<sup>2</sup> Paper presented on the 9<sup>th</sup> Professional Meeting "Shared cataloguing system" organized by the Serbian Academic Library Association and held in the National Library of Serbia, Belgrade, 10 -11 October 2003.

терна организација елемената који чине документ представља његову структуру. Поједини елементи документа могу припадати различитим типовима медија, као што су текст, слике, видео и аудио записи. Претраживање мултимедијалних докумената подразумева проналажење оних докумената у посматраној колекцији чији садржај одговара датом критеријуму. У мултимедијалном амбијенту претраживање по садржају односи се како на текст, тако и на друге типове медија (статичне слике, видео записи).

Резултати у области проналажења мултимедијалних информација присутни су више од деценије. До данас је развијен одређени број система који омогућавају претраживање мултимедијалних докумената. У овом тексту биће приказана четири репрезентативна система.

Један од првих система за претраживање мултимедијалних докумената, MINOS [1], конципиран је тако да рукује документима фиксне структуре. Документи MINOS-а могу да садрже текст, слике, звучне записе и додатне атрибуте (структуриране метаподатке). Претраживање колекције оваквих докумената има следеће могућности: (1) претраживање по садржају текстуалних елемената, (2) претраживање по садржају структурираних метаподатака, (3) претраживање по егзистенцији слике или звучног записа у документу. Претраживање по садржају мултимедијалних докумената MINOS-а дакле, није потпуно: није могуће претраживати документе по садржају њихових слика односно звучних записа.

Систем MULTOS [2] је, по својим карактеристикама, сличан MINOS-у. Основну разлику између ова два система представља могућност MULTOS-а да рукује документима различите структуре. Претраживање по садржају појединих елемената документа функционално је еквивалентно MINOS-у: слике и звучни записи не могу се претраживати по садржају већ само по егзистенцији у документу.

Модел за проналажење мултимедијалних докумената приказан у [3] користи концепте предикатске логике за репрезентацију садржаја докумената и његових елемената. Овај модел омогућава руковање документима различите структуре. Претраживање докумената могуће је како по карактеристикама њихове структуре, тако и по садржају појединих елемената. Елементи документа могу припадати различитим типовима медија. Да би претраживање по елементима одређеног типа медија било могуће, потребно је изразити садржај датог елемента концептима предикатске логике коју користи модел. На овај начин, модел је могуће проширити подршком за различите типове

Particular document elements can belong to different media types, such as text, images, video, and audio clips. Multimedia document retrieval represents the retrieval of documents contained in a given document collection that conform to the given search criterion. The multimedia document environment assumes content-based retrieval (CBR) for all media types (eg., text retrieval, retrieval by image content, retrieval by video content).

Research results in the field of multimedia information retrieval exist for more than a decade. A number of systems providing multimedia document retrieval has been developed. This paper discusses features of the four representative systems.

One of the first multimedia document retrieval systems, MINOS [1], handles documents with fixed structure. MINOS documents can contain text, images, audio recordings and attributes (alphanumerical metadata). The retrieval of documents in a collection has the following features: (1) content-based retrieval on textual elements, (2) CBR on the content of attributes, and (3) retrieval based on the existence of an image or an audio clip in a document. Therefore, the retrieval of documents based on their content is not complete in MINOS – it is not possible to search document collections based on the content of their images or audio clips.

The MULTOS system [2] is functionally similar to MINOS. The main difference represents the ability of MULTOS to handle documents with different structure. The retrieval functionality is equivalent to that of MINOS – images and audio clips can be queried only on their existence, rather than their content.

The multimedia information retrieval model presented in [3] uses concepts of predicate logic for representing content. This model is able to handle documents with different structure. Document retrieval uses information on both document structure and element content. Document elements can belong to different media types. In order to support retrieval using a certain media type, a model of content representation using predicate logic must be formulated. Thus the retrieval model is extensible with support for different media types, provided that their content representation uses the concepts of predicate

медија уз услов да се репрезентација њиховог садржаја изражава датим концептима. Наметање сопственог модела репрезентације садржаја је и основно ограничење овог модела проналажења докумената.

У циљу превазилажења недостатака постојећих система за IR реализован је Прошириви систем за проналажење мултимедијалних докумената (*Extensible Multimedia Information Retrieval System*, XMIRS), детаљно приказан у дисертацији [4]. Полазна тачка његовог развоја била је да се омогући проширивост система. Проширивост се, у овом случају, огледа у два аспекта: (1) проширивост система подршком за претраживање различитих типова медија по њиховом садржају и (2) проширивост система различитим моделима проналажења докумената. Формална спецификација софтверске архитектуре XMIRS-а и модела проналажења докумената који су намењени за XMIRS дата је у [4, 5].

Имплементација мрежне дигиталне библиотеке докторских и магистарских радова представља наставак развоја Библиотечког информационог система БИСИС [18] намењена раду са оригиналним документима у дигиталном облику. У наставку текста дат је преглед основних карактеристика архитектуре XMIRS-а и илустративни примери његове примене у Мрежној дигиталној библиотеци који приказују комбиновање претраге по текстуалним елементима докумената, сликама и видео записима.

## 2. Архитектура XMIRS-а

Систем XMIRS рукује структурираним мултимедијалним документима чији се садржај представља XML језиком [6]. XML је пре свега намењен репрезентовању садржаја структурираних текстуалних докумената али се може употребити и за опис садржаја мултимедијалних докумената. Структура XML докумената је хијерархијска, заснована на структури стабла. Сложеније структуре докумената могуће је формирати коришћењем линкова које дефинише XLink стандард [7].

Структура докумената којима рукује XMIRS описује се XML Schema језиком [8, 9]. XMIRS може да рукује документима различитих типова. За сваки тип документа потребно је дефинисати његову XML Schema спецификацију.

Употреба стандардног HTTP протокола [10] за испоруку докумената омогућава широку доступност докумената различитим типовима клијената у web окружењу. URI стандард [11] за јединствену идентификацију ресурса обезбеђује независност апликација од начина испоруке докумената и локације њиховог складиштења.

logic. Imposing its own model of representing content is the main drawback of this retrieval model.

In order to overcome the shortcomings present in existing multimedia IR systems and models an Extensible Multimedia Information Retrieval System (XMIRS), presented in detail in [4], has been developed. The main focus of XMIRS development was to provide system extensibility. System extensibility comes in two ways: (1) extensibility by supporting multiple media types in document retrieval and (2) extensibility with multiple document retrieval models. Formal specification of XMIRS software architecture and retrieval models suitable for XMIRS is given in [4, 5].

The implementation of the Networked Digital Library of Theses and Dissertations (NDLTD) is a result of continuous development of the BISIS library information system [18] that is able to handle original documents in their digital form. The rest of the paper gives an overview of main characteristics of XMIRS architecture and examples of applying XMIRS to the Networked Digital Library that illustrate document retrieval combining search on textual elements, images, and video clips.

## 2. XMIRS Architecture

The XMIRS handles structured multimedia documents expressed by the XML language [6]. XML is primarily a language for representing structured textual documents, but can be seamlessly used to represent multimedia documents. The structure of XML documents is a hierarchical, tree-based structure. More complex document structures can be formed using links defined by the XLink standard [7].

The structure of documents handled by XMIRS is defined by XML Schema [8, 9] documents. XMIRS is able to handle documents of different types. Each document type must be defined by its XML Schema specification.

The use of the standard HTTP protocol [10] for document delivery provides wide document availability to different types of clients in a web environment. The URI standard [11] for unique resource identification is used to provide applications with independence of document delivery methods and document storage locations. Document types handled

Типови докумената којима рукује XMIRS, описани својим XML Schema спецификацијама, доступни су на адресама формираним по следећем шаблону: `<protocol>://<host>[:<port>]/schemas/<schema-name>`

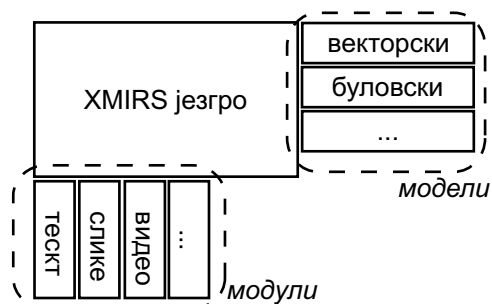
Адреса документа (или његовог сегмента одређеног XPointer синтаксом) је следећег формата:

`<protocol>://<host>[:<port>]/docs/<schema-name>/<doc-id>[#<xpointer-expr>]`

XMIRS систем може да рукује са више колекција докумената. Под *колекцијом докумената* подразумева се онај скуп докумената који се претражује приликом процеса проналажења. Два аспекта проширивости система се, са становишта његове архитектуре, огледају у постојању два типа проширења система. *Модули* XMIRS-а, као један тип проширења, намењени су претраживању елемената документа одређеног типа медија (текст, слике, видео, итд). Основне функције модула су:

- (1) припрема садржаја појединих елемената документа за претраживање (тзв. индексирање) и
- (2) претраживање садржаја индексираних елемената.

Један модул XMIRS-а, према томе, обезбеђује парцијално претраживање колекције докумената по једном типу медија односно једном скупу елемената документа. Парцијални резултати претраге се могу комбиновати на различите начине за потребе израчунавања коначног резултата. *Модел* XMIRS-а, као други тип проширења, намењени су израчунавању коначног резултата претраживања докумената на основу парцијалних резултата претраге које обезбеђују модули. **Слика 1** приказује основне елементе архитектуре XMIRS-а која се састоји од језгра, намењеног повезивању проширења система и комуникацији са корисником, и модула и модела система.



Слика 1. Архитектура XMIRS система

Комуникација XMIRS-а са корисником одвија се путем XML докумената. Комуникација обухвата упите упућене систему и резултате претраживања. Синтакса упита и резултата одређена је изабраним моделом проналажења докумената. Упити упућени систему састоје се из више елемен-

by XMIRS, specified by corresponding XML Schema documents, are available at addresses conforming to the following pattern:

`<protocol>://<host>[:<port>]/schemas/<schema-name>`

Address of a document (or its part identified by an XPointer expression) conforms to the following pattern:

`<protocol>://<host>[:<port>]/docs/<schema-name>/<doc-id>[#<xpointer-expr>]`

XMIRS is able to handle multiple document collections. A *collection* is a set of documents used in a retrieval session. The two aspects of system extensibility are implemented as two types of system extensions. The first type, *modules*, provide retrieval functionality for a particular media type (text, images, video, etc). The basic functions of a module are the following:

- (1) forming the representation of element content used in retrieval, ie. indexing, and
- (2) retrieval of indexed elements.

Each XMIRS module provides partial retrieval functionality on a single media type or a single subset of document elements. Partial retrieval results can be combined in different ways to calculate the final retrieval result. XMIRS *models*, as the other type of system extensions, deal with the final result computation using partial retrieval results provided by retrieval modules. **Figure 1** illustrates the main elements of XMIRS architecture comprising the system core, used for integrating extensions and communication with users, modules, and models.

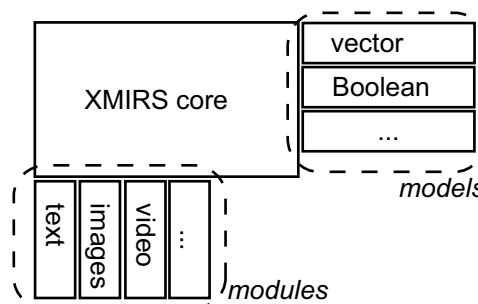


Figure 1. XMIRS system architecture

Communication between users and the XMIRS system is based on XML documents and comprises queries posted to the system and retrieval results. Query and result syntax is defined by the chosen retrieval model. Queries contain multiple elementary queries (as subdocuments of the query document).

тарних упита (поддокумената упитног документа). Израчунавање резултата елементарних упита задат је појединих модула за претраживање.

Проширивост XMIRS-а модулима за рад са различитим типовима медија и различитим моделима проналажења докумената омогућава прилагођавање система конкретним потребама појединих информационих система. XMIRS је конфигурабилан систем – његовом конфигурацијом обухваћене су дефиниције типова докумената, параметри складиштења докумената, као и параметри појединих модула и модела. Поред тога, као модуле система могуће је користити постојећа софтверска решења за претраживање појединих типова медија. Примена XMIRS-а на реалном систему дигиталне библиотеке приказана је у наредном одељку.

### 3. Примена XMIRS-а у Мрежној дигиталној библиотеци

Пројекат мреже дигиталних библиотека докторских и магистарских теза (*Networked Digital Library of Theses and Dissertations*, NDLTD) [12] започет је на универзитету Virginia Tech, САД. Главни циљ NDLTD пројекта је једноставна и ефикасна централизована каталогизација докторских и магистарских теза одбрањених у оквиру институција-чланова. NDLTD систем се састоји од чворова у мрежи који представљају аутономна складишта докумената. Оригинални, комплетни документи се чувају на одговарајућем чвору у неком од погодних формата дигиталних докумената (нпр. PDF). Чворови мреже омогућавају проналажење докумената које сами складиште. Процес проналажења докумената користи метаподатке о документима који чине саставни део описа докумената. У формирању метаподатака за документ учествују аутор документа и надлежни библиотекари. Приликом прегледа резултата проналажења, могуће је преузети и оригинални облик траженог документа. NDLTD мрежа не намеће посебне захтеве пред функционалност проналажења докумената имплементираних у оквиру појединих чворова. Централни систем за проналажење докумената [13] постоји као посебан чвор мреже који омогућава проналажење докумената складиштених на свим осталим чворовима. Проналажење користи метаподатке формиране на оригиналним чворовима. Централни систем не складишти оригиналне документе, већ само URL адресе до оригиналног чвора на коме су складиштени. Имплементација NDLTD чвора Универзитета у Новом Саду приказана је у [14].

У [4] приказан је предлог проширене структуре NDLTD докумената који, поред оригиналног електронског документа и метаподатака који

The calculation of particular elementary queries is a task of specific retrieval modules.

Extensibility of XMIRS with support for different media types and different retrieval models enables system customization in order to meet the specific needs of a particular application. XMIRS is a configurable system; its configuration contains document type definitions, document storage parameters, as well as parameters of particular modules and models. Besides, existing software packages can be integrated into XMIRS in the form of modules. The next section presents the application of XMIRS to a real system of Networked Digital Library.

### 3. Applying XMIRS to Networked Digital Library

The Networked Digital Library of Theses and Dissertations project originates from the Virginia Tech University, USA. The main goal of the NDLTD project is a simple and efficient centralized cataloguing of master and PhD theses defended and published in member institutions. The NDLTD system is a network of nodes (institutions) representing autonomous document storages. Original electronic documents are stored at the corresponding node in a portable electronic format such as PDF. Each node in the network provides document retrieval for the collection it stores. The document retrieval process uses document metadata. The metadata is stored along with documents. The creation of metadata for the given document is a task carried out by the document author and the librarian. Retrieval results consist of document metadata giving a brief overview of the original document content. However, the retrieval results contain links to the original document themselves. The NDLTD network does not impose any particular retrieval functionality to its member nodes. The NDLTD Union Catalog [13] is a special node in the network that provides unified retrieval of documents stored on all other nodes. This retrieval uses the metadata originating from other nodes. The central catalog, however, stores only URL addresses to documents stored on original nodes, rather than storing all documents by itself. The implementation of an NDLTD node used at the University of Novi Sad is presented in [14].

The XMIRS case study in [4] presents a proposition for extending the structure of NDLTD documents to accommodate, besides alphanumeric

описују његов садржај, поседује и слике и видео записе као посебну врсту метаподатака којима се ближе описује садржај документа. Слика 2 приказује ову структуру.

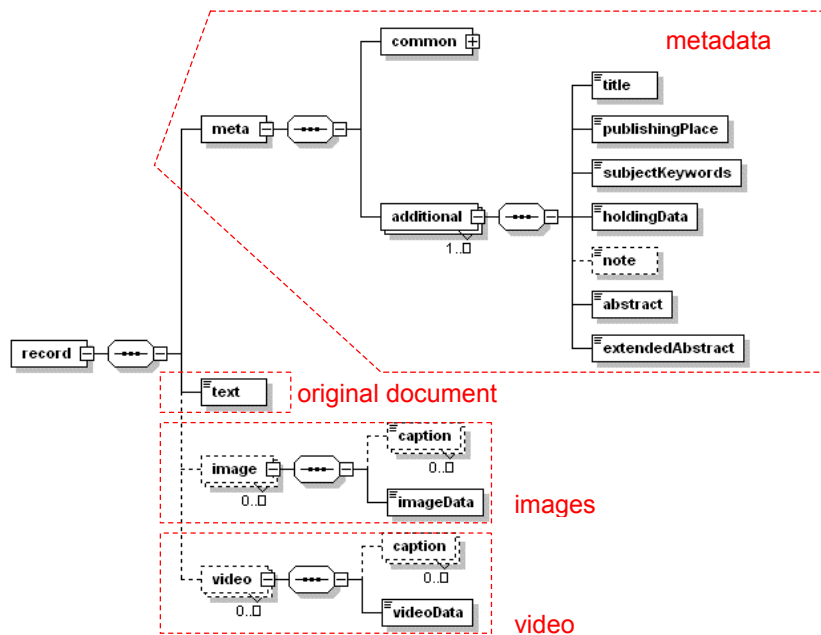
Метаподаци чији садржај је нумеричког типа или припада контролисаном речнику смештени су у оквиру подстабла чији је корен елемент *common*. Другу групу метаподатака, са коренским елементом *additional*, чине елементи чији се текстуални садржај може изразити на више језика (нпр. апстракт). За претраживање прве групе метаподатака може се користити језик XPath [15] који подразумева егзактно поређење садржаја метаподатака са критеријумом претраге. Софтверски систем који имплементира претраживање колекција XML докумената помоћу XPath упита је Apache Xindice [16]. Xindice је, у оквиру XMIRS-а, употребљен у форми модула намењеног за претраживање текстуалних елемената докумената помоћу XPath језика.

Претраживање друге групе метаподатака захтева могућности које се срећу у системима за претраживање неструктурираног текста (*full text*): логичке, тежинске и близинске операторе, употреба цокер-знакова, претраживање по корену речи и слично. Такву функционалност имплементира Apache Lucene библиотека [17] за претраживање неструктурираних текстуалних докумената. Ова библиотека је интегрисана у XMIRS у виду његовог модула. Lucene се користи и за *full text* претраживање оригиналног документа датог у PDF формату.

metadata and the original documents, images and video clips as a new form of metadata representing document content. Figure 2 presents the new document structure.

Numerical metadata and metadata that take values from a controlled vocabulary is stored in a document subtree with the element *common* as its root. The other part of metadata, stored as a subtree rooted by the element *additional*, are textual elements whose contents can be expressed in multiple natural languages (such as document abstract). Retrieval of the first part of metadata is carried out using the XPath query language [15] that is based on exact matching between the content (metadata) and the query. The software system that implements searching document collections using XPath queries is Apache Xindice [16]. Xindice is integrated into XMIRS in the form of an XMIRS module providing retrieval on textual document elements by means of XPath queries.

The retrieval of the other group of metadata needs the full-text retrieval functionality common to many systems handling unstructured textual documents: the use of logical and proximity operators, term weighting, wildcard characters, word stemming, etc. Such functionality is implemented in the Apache Lucene software library [17] for retrieval of unstructured textual documents. This library is integrated into XMIRS in the form of an XMIRS module. Lucene is used for full-text searching on the original PDF document as well.



Слика 2. Структура проширених NDLTD докумената  
Figure 2. The extended structure of NDLTD documents

Претраживање докумената помоћу распрострањеног Dialog упитног језика имплементирано је у

Searching document collections using the Dialog query language is implemented in the BISIS

Библиотечком информационом систему БИСИС. Текст сервер БИСИС-а [18] је интегрисан у систем у форми модула и омогућава претраживање докумената помоћу концепта префикса и синтаксе Dialog језика.

Претраживање слика по садржају имплементирано је у оквиру Oracle9i базе података, односно њеног interMedia додатка [19]. Претраживање слика по садржају користи слику-узорак као упит са којом се пореде све слике у колекцији. Поређење слика по садржају користи више критеријума сличности слика:

- (1) сличност по глобалном присуству појединих боја у слици,
- (2) сличност по распореду региона обојених поједином бојом,
- (3) сличност по текстури слике,
- (4) сличност облика приказаних у слици и
- (5) сличност по просторном распореду објекта у слици.

Упит за претраживање слика се, према томе, састоји од слике-узорка и скупа нумеричких параметара који дефинишу релевантност појединих елемената сличности приликом израчунавања резултата. XMIRS систем је за ову прилику допуњен модулом који представља везу змеђу XMIRS-а и Oracle9i базе података. Слике се, поред тога, могу претраживати и по придруженим текстуалним описима помоћу Lucene модула.

Претраживање видео записа подразумева претходну анализу садржаја видео записа и репрезентацију садржаја путем кључних фрејмова. Кључни фрејмови представљају статичне слике, најчешће директно преузете из видео записа, које репрезентују визуелни садржај једног кадра садржаног у видео запису. На тај начин се садржај једног видео записа састављеног из низа кадрова може представити низом статичних слика. Овај низ статичних слика се може претраживати по садржају на исти начин како је то урађено у претходном случају, употребом Oracle9i модула. Претходна анализа садржаја видео записа и екстракција кључних фрејмова је задатак посебног софтверског система, у овом случају IBM CueVideo [20]. Овај систем је интегрисан у XMIRS у форми посебног модула.

#### 4. Примери претраживања помоћу XMIRS-а

**Први пример** употребе XMIRS-а за претраживање проширених NDLTD докумената обухвата два различита текстуална упита. У овом примеру потребно је пронаћи све документе писане на енглеском језику, у чијем проширеном апстракт се јавља фраза „information systems“. Први део критеријума претраге обухвата претраживање прве

library information system. The BISIS text server [18] is integrated into XMIRS in the form of an XMIRS module and provides retrieval on document metadata using the Dialog query language.

Content-based retrieval of images is implemented in Oracle9i database, specifically its interMedia cartridge [19]. Retrieving images based on content uses a sample image as a query that is compared to all images in a given collection. Image comparison in Oracle9i uses multiple comparison criteria:

- (1) similarity by global presence of colors,
- (2) similarity by spatial position of colored regions,
- (3) similarity of image textures,
- (4) similarity of shapes presented by in the image, and
- (5) similarity by spatial relations of objects in the image.

Hence, query for image retrieval consists of a sample image and a set of numerical parameters determining the relevance of particular image similarity criteria. XMIRS system is supplied with a module acting as a mediator between XMIRS and the Oracle9i database. Images can be retrieved by textual annotations associated with the images by Lucene module as well.

The retrieval of video clips assumes prior analysis of the clips' contents. The content of video clips is usually represented by means of keyframes. Keyframes are static images, usually directly extracted from the video clip, that represent visual content of a single shot contained in the video clip. This way contents of a video clip can be substituted by a series of static images representing its content. Retrieval of video clips can then utilize techniques developed for static images, as implemented using the Oracle9i module. Prior analysis of video clip contents and keyframe extraction is a task of the special software package, IBM CueVideo [20]. This software package is integrated into XMIRS in the form of an XMIRS module.

#### 4. Examples of document retrieval using XMIRS

**The first example** of applying XMIRS to the retrieval of NDLTD documents comprises two different textual queries. This example searches for all documents written in English whose extended abstract contains the phrase “information systems”. The first part of the search criterion deals with searching the first group of metadata, in this case the element

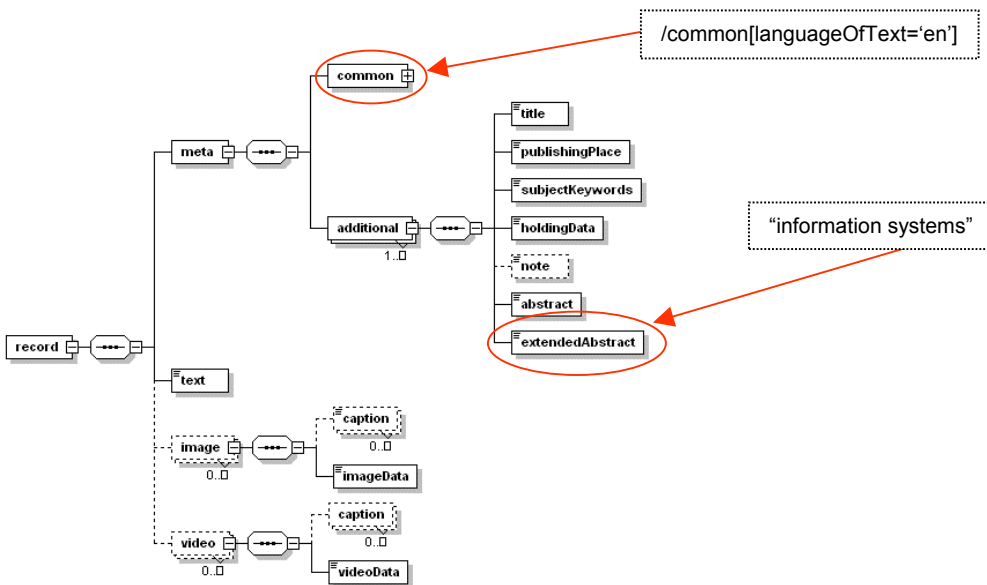
групе метаподатака, у овом случају елемента *languageOfText*, чији садржај припада контролисаном речнику. За тај део претраге користи се Xindice модул, са упитом формираним према XPath синтакси. У овом случају XPath упит гласи: `/common[languageOfText='eng']`. Претраживање проширених апстрактата обавља Lucene модул код кога се фраза која се тражи смешта између двоструких наводника, тако да одговарајући упит гласи: **"information systems"**. Укупни упит обухвата претходна два елементарна упита и, према векторском моделу проналажења докумената изражава се следећим XML документом:

```
<?xml version="1.0" encoding="UTF-8"?>
<vectorQuery>
  <elemQueries>
    <elemQuery index="common.metadata">
      <xindice>/common[languageOfText='eng']</xindice>
    </elemQuery>
    <elemQuery index="ex.abstract">
      <lucene>"information systems"</lucene>
    </elemQuery>
  </elemQueries>
  <metric>L_inf</metric>
  <maxHits>-1</maxHits>
</vectorQuery>
```

Критеријум претраге из овог примера и елементи документа на које се односи приказани су на слици 3.

*languageOfText*. This part of the search uses the Xindice module and the elementary query expressed in XPath syntax. The corresponding elementary query is `/common[languageOfText='eng']`. Searching extended abstracts is performed by the Lucene module, where the phrase searched for is put between double quotes. Therefore, the Lucene query is: **"information systems"**. The overall query contains both elementary queries and, according to the vector model of document retrieval, is expressed as the following XML document:

The search criterion in the first example and document elements used in retrieval are presented in Figure 3.



Слика 3. Пример комбиновања претраге по метаподацима  
Figure 3. An example of combining different metadata retrieval

**Други пример** представља комбиновано претраживање докумената по апстрактима и придруженим сликама. Потребно је пронаћи документе у чијем апстракту се јавља фраза „lung cancer“, а придружена слика личи на дату слику-узорак. За први део критеријума претраге користи се Lucene

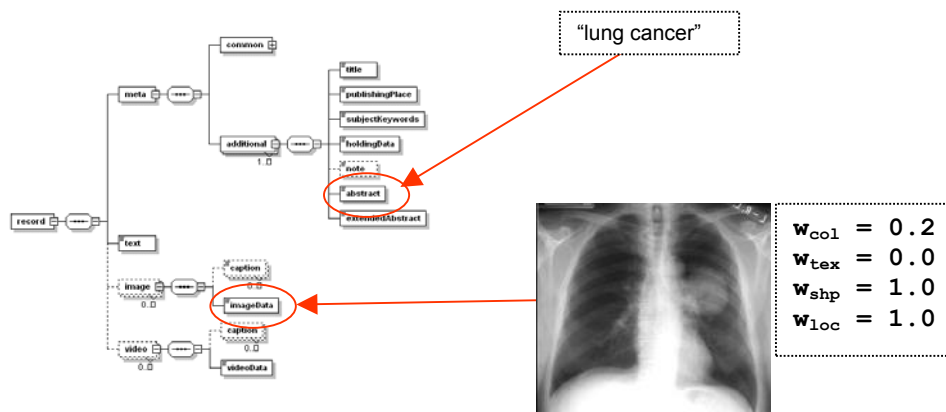
**The second example** illustrates combined search on abstracts and associated images. This example searches for documents whose abstract contains the phrase “lung cancer”, and the associated image is similar to the given image sample. The first part of the search criterion is handled by the Lucene



модул за претраживање по апстрактима, док се за други део критеријума користи Oracle9i модул за претраживање по придруженим сликама. У складу са синтаксом Lucene модула први део критеријума гласи "lung cancer". Други део критеријума претраге састоји се из слике-узорка (рентгенског снимка плућа) и скупа нумеричких параметара који дефинишу критеријум сличности слика са узорком. У овом случају, сличност слика по боји је од мале важности, сличност по текстури није од значаја, а сличност по приказаним облицима и њиховом положају у простору максимална, што се изражава одговарајућим нумеричким вредностима. На слици 4 приказан је критеријум претраге из овог примера и елементи докумената на које се критеријум односи. Укупан упит се, према векторском моделу проналажења докумената, изражава следећим XML документом (садржај елемента *image* који представља Base64-кодиран садржај слике је изостављен):

module, while the other part is handled by the Oracle9i database. The first elementary query is, according to Lucene syntax, "lung cancer". The second elementary query is comprised of an image sample (an X-ray image) and a set of numerical parameters defining the image similarity criteria. In this case, similarity by color is less important, texture similarity is completely ignored, and similarity of shapes and their spatial constraints is of maximum importance. Figure 4 presents the search criterion of the second example and document elements used in retrieval. The overall query, using the means of the vector model, is expressed by the following XML document (the Base64-encoded contents of the *image* element are left out):

```
<?xml version="1.0" encoding="UTF-8"?>
<vectorQuery>
  <elemQueries>
    <elemQuery index="abstract">
      <.lucene>&apos;lung cancer&apos;</.lucene>
    </elemQuery>
    <elemQuery index="images">
      <intermedia>
        <image contentType="image/jpeg">...</image>
        <weights color="0.2" texture="0.0" shape="1.0" location="1.0"/>
      </intermedia>
    </elemQuery>
  </elemQueries>
  <metric>L_2</metric>
  <maxHits>50</maxHits>
</vectorQuery>
```



Слика 4. Пример комбиновања претраге по текстуалним метаподацима и сликама  
Figure 4. Combining retrieval on textual metadata and images

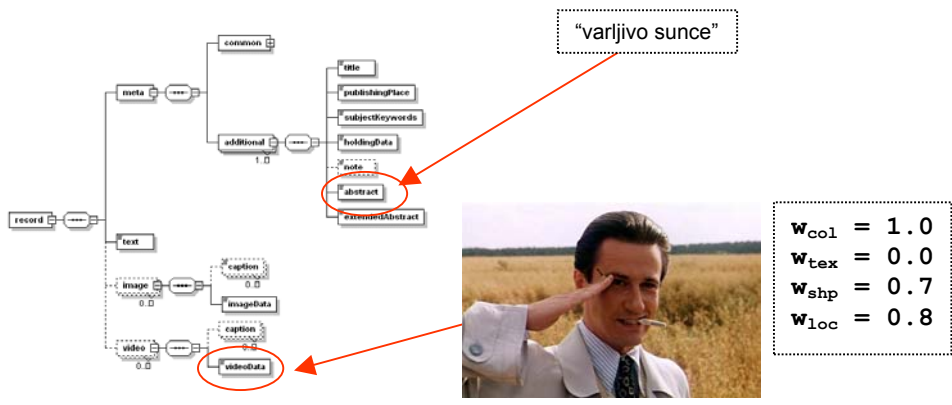
**Трећи пример** илуструје могућност употребе логичких оператора приликом комбиновања упита. Потребно је пронаћи документе у чијем апстрактку се јавља фраза „varljivo sunce“ **или** при-

**The third example** illustrates the use of logical operators in forming complex queries. The task is to retrieve documents whose abstract contains the phrase “varljivo sunce” **or** the associated video clip contains

дружени видео запис садржи кадар који је сличан датој слици. Први део упита се, слично претходним примерима, формулише као **"varljivo sunce"** и упућује Lucene модулу. Други део упита се формулише као критеријум за претраживање по сликама који се упућује Oracle9i модулу, али се он сада односи на слике које су добијене анализом видео записа и екстракцијом кључних фрејмова које је обавио CueVideo модул. У овом случају сличност слика по боји је од максималне важности, док је сличност по облицима приказаним сликом и њиховом просторном распореду мање значајна. Текстура слике није од значаја. Слика 5 илуструје критеријум претраге из овог примера. Укупан упит се, према проширеном буловском моделу, изражава следећим XML документом:

a shot visually similar to the given sample image. The first elementary query, similarly to previous examples, is expressed as **"varljivo sunce"** and passed to the Lucene module. The other elementary query is an image retrieval query passed to the Oracle9i module. However, in this case the images retrieved by the Oracle9i module are those generated by the CueVideo module as a result of video clip analysis and keyframe extraction. Image similarity by color is of the highest importance, while the similarity by shape and spatial relations is less significant. Image texture is of no importance. Figure 5 illustrates the search criterion of the third example. The overall query is, according to the extended boolean retrieval model, expressed by the following XML document:

```
<?xml version="1.0" encoding="UTF-8"?>
<boolQuery>
  <operator type="OR" metric="L_2" maxHits="100">
    <elemQuery index="abstracts">
      <.lucene>&quot;varljivo sunce&quot;</.lucene>
    </elemQuery>
    <elemQuery index="videos">
      <intermedia>
        <image type="image/jpeg">...</image>
        <weights color="1.0" texture="0.0" shape="0.7" location="0.8"/>
      </intermedia>
    </elemQuery>
  </operator>
</boolQuery>
```



Слика 5. Пример комбиновања претраге по текстуалним метаподацима и видео записима  
 Figure 5. Combining retrieval on textual metadata and video clips

### 5. Закључак

Системи за претраживање колекција мултимедијалних докумената би требало да омогуће претраживање докумената по њиховом садржају без обзира на тип медија коме поједини елементи документа припадају. XMIRS систем је конципиран тако да обезбеди проширивост у смислу додавања подршке за претраживање по различитим типовима медија и израчунавање резултата претраге по различитим моделима проналажења докумената. У том смислу, XMIRS превазилази

### 5. Conclusions

Multimedia document retrieval systems should facilitate content-based retrieval using any media type present in documents. XMIRS system is designed to provide extensibility by adding support for retrieval of different media types and final result calculation using different document retrieval models. This way XMIRS overcomes the main shortcomings of existing multimedia document retrieval systems.

ограничења постојећих система за претраживање мултимедијалних докумената.

Конфигурабилност и проширивост XMIRS-а омогућава његову употребу у различитим применама. У овом раду је приказана примена XMIRS-а у систему Мрежне дигиталне библиотеке магистарских и докторских радова (NDLTD). Дат је предлог проширења структуре докумената којима рукује NDLTD систем мултимедијалним елементима (сликама и видео записима) и приказано решење које омогућава комбиновано претраживање по разнородним елементима докумената.

Решење претраживања мултимедијалних докумената које нуди XMIRS може се употребити као основа за изградњу нове верзије система БИСИС која представља мултимедијални информациони систем. Функције оваквог система обухватају, поред руковања класичним UNIMARC репрезентацијама докумената, и руковање оригиналним документима у електронском облику, као и претраживање фонда коришћењем различитих типова медија.

XMIRS extensibility and configurability facilitates its use in different applications. This paper has presented the application of XMIRS to the Networked Digital Library of Theses and Dissertations. A proposition for extending the current NDLTD document structure with multimedia elements (images and video clips) is given. A solution for combined multimedia retrieval using different media types is described.

This solution for multimedia document retrieval may serve as a basis for development of a new version of the BISIS library information system, that represents a multimedia information system. Functions of such a system include, besides handling of classical UNIMARC document representations, handling of original documents in electronic form, and retrieval of documents using different media types as well.

## References

- [1] S. Christodoulakis, M. Theodoridou, F. Ho, M. Papa, and A. Pathria. Multimedia Document Presentation, Information Extraction, and Document Formation in MINOS: A Model and a System. *ACM Transactions on Office Information Systems*, 4(4):345-383, 1986.
- [2] E. Bertino, F. Rabitti, and S. Gibbs. Query Processing in a Multimedia Document System. *ACM Transactions on Office Information Systems*, 6(1):1-41, 1988.
- [3] C. Meghini, F. Sebastiani, and U. Straccia. A Model of Multimedia Information Retrieval. *Journal of the ACM*, 48(5):909-970, 2001.
- [4] B. Milosavljević. *Proširivi sistem za pronalaženje multimedijalnih dokumenata*. Doktorska disertacija, Fakultet tehničkih nauka Novi Sad, 2003. <http://diglib.ns.ac.yu/ndltd/docs/set1/ndltd143/xmirs.pdf>
- [5] B. Milosavljević and Z. Konjović. Design of an XML-Based Extensible Multimedia Information Retrieval System. In *Proceedings of IEEE Multimedia Software Engineering 2002*, pp. 114-121, Newport Beach, CA, 2002.
- [6] T. Bray, J. Paoli, C. M. Sperber-McQueen, and E. Maler. *Extensible Markup Language (XML) 1.0 (Second Edition)*. W3C Recommendation, 2000. <http://www.w3.org/TR/REC-xml>
- [7] S. DeRose, E. Maler, and D. Orchard. *XML Linking Language (XLink)*. W3C Recommendation, 2001. <http://www.w3.org/TR/xlink>
- [8] H. S. Thompson, D. Beech, M. Maloney, N. Mendelsohn. *XML Schema Part 1: Structures*. W3C Recommendation, 2001. <http://www.w3.org/TR/xmlschema-1>
- [9] P. V. Biron and A. Malhotra. *XML Schema Part 2: Datatypes*. W3C Recommendation, 2001. <http://www.w3.org/TR/xmlschema-2>
- [10] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee. *Hypertext Transfer Protocol – HTTP/1.1*. IETF RFC 2616. <http://www.ietf.org/rfc/rfc2616.txt>
- [11] T. Berners-Lee, R. Fielding, and L. Masinter. *Uniform Resource Identifiers (URI): Generic Syntax*. IETF RFC 2396. <http://www.ietf.org/rfc/rfc2396.txt>
- [12] *Networked Digital Library of Theses and Dissertations*. Virginia Tech, USA. <http://www.ndltd.org>
- [13] *NDLTD Union Catalog*. VTLS Inc., USA. <http://www.vtls.com/ndltd>

- [14] Д. Сурла, З. Коњовић, Б. Милосављевић, Г. Сладић, З. Протић, С. Комазец, Д. Окановић. Приказ реализације мрежне дигиталне библиотеке докторских, магистарских и дипломских радова. *Инфотека: часопис за информатику и библиотекарство*, 5(2004)1-2, стр. 75-86.  
D. Surla, Z. Konjović, B. Milosavljević, G. Sladić, Z. Protić, S. Komazec, and D. Okanović. An Overview of the Implementation of the Networked Digital Library of Theses and Dissertations. *Infotheca: Journal of Informatics and Librarianship*, 5(2004)1-2, pp. 75-86.
- [15] J. Clark and S. DeRose. *XML Path Language (XPath) Version 1.0*. W3C Recommendation, 1999.  
<http://www.w3.org/TR/xpath>
- [16] *Apache Xindice*. The Apache Software Foundation. <http://xml.apache.org/xindice>
- [17] *Apache Lucene*. The Apache Software Foundation. <http://jakarta.apache.org/lucene>
- [18] Б. Милосављевић. *Текст сервер UNIMARC записа*. Магистарска теза, Факултет техничких наука, Нови Сад 1999. <http://diglib.ns.ac.yu/ndltd/docs/set1/ndltd6/textsrv.pdf>  
B. Milosavljević. Text Server For UNIMARC Records. Master's thesis, Faculty of Engineering, Novi Sad, 1999. <http://diglib.ns.ac.yu/ndltd/docs/set1/ndltd6/textsrv.pdf> (in Serbian)
- [19] *Oracle9i Database*. Oracle Corporation. <http://www.oracle.com/ip/dep/otn/database/oracle9i>
- [20] *CueVideo*. IBM Almaden Research Center. <http://www.almaden.ibm.com/projects/cuevideo.shtml>