ICADL 2007, Hanoi, 11 December 2007 Session 2C: European DLs I

Invited Talk on

Information Access through Digital Library Systems

Maristella Agosti

Information Management Systems (IMS) Research Group

Department of Information Engineering (DEI)

University of Padua, Italy



Università degli Studi di Padova







Introduction: University of Padua, DEI, IMS						
Information Access: the Three Main Actors	Final Aim of a DL/DLS from					
Information Access System	ine point of view of the user					
O Centralized Systems	Architectures of DL Systems					
O Centralized Systems – Distributed Access	Available before International					
O Distributed Access also Through the Web	Programs and Conferences					
DELOS View of DL, DLS and DLMS						
DELOS Reference Model Overview DELOS Achievements						
Service Oriented Architecture						
DL Service: the Case of Annotations	Contribution of the					
Evaluation of a DLMS	University of Padua to					
Personalization through Web Log Analysis	DELOS Achievements					



- Introduction: University of Padua, DEI, IMS
- Information Access: the Three Main Actors
- Information Access System
 - Centralized Systems
 - O Centralized Systems Distributed Access
 - Distributed Access also Through the Web
- DELOS View of DL, DLS and DLMS
- DELOS Reference Model Overview
- Service Oriented Architecture
- DL Service: the Case of Annotations
- Evaluation of a DLMS
- Personalization through Web Log Analysis



University of Padua, Italy

University of Padua:

- 1222: first academic year, second oldest university in Italy
- □ Organization:
 - O Faculties: teaching (13)
 - O Departments: research (more than 60)

□ Staff:

- O 2500 research/teaching
- O 2000 technical and administration

□ Students:

- ◯ 57000/60000
- O More than 6000 graduate each year

http://www.unipd.it/





University of Padua: History



Galileo Galilei's desk



Astronomical Observatory



Anatomy Theatre: built in 1594 first permanent anatomy theatre in the world.



Main building of Padua University "Palazzo del Bo"



First woman graduate in the World: Elena Lucrezia Piscopia (degree in Philosophy in 1678)



Botanical Gardens founded in 1545



Department of Information Engineering - DEI

DEI:

Established in 1987

Research:

- **O** Applied Optics
- **O** Automatics
- O Bioengineering
- O Computer Science
- **O** Electronics
- **O** Telecommunications

Staff:

- 100 research/teaching
- O 35 technical and administration
- 100 PhD, research grant holders





http://www.dei.unipd.it/

UNIVERSITA DI PADOVA - Dipartimento di Ingegneria dell'Informazione - Mazilla Firefox De dodda (walkza) (ye Segnibli (knownada ikinden/Ebesone - 1) (v) (v) (v) (v) (v) (v) (v) (v) (v) (v		
Come iniziare		
HOME PERSONE D	IPARTIMENTO RICERCA SCIENTIFICA DIDATTICA DOTTORATO NEWS Cerca	enç
-> HELPDESK <	UNIVERSITA' DI PADOVA - Dipartimento di Ingegneria dell'Informazione	Accedi aree ris
UNIVERSITA' DI PADOVA - D Ele Modifica Youalizza Va Come intere Si Ultime notizie HOME PERSONE -> HELPDESK <- LATEST NEWS UNI EN 150 9001;2000 CISO	Colloquia Seminari riguardanti le attivita' di studio e ricerca del Dipartimento. Se non diversamente speci seminari si svolgono alle ore 15 in Aula	ficato, i password:
	Biblioteca Interdipartimentale di Ingegneria dell'Informazione e Ingegneria Elettric Biblioteca Interdipartimentaledi Ingegneria dell'Informazione e Ingegneria Elettrica via Giovann Gradenigo 6A 35131	a leggi
	Distinguished Lecturer Series	
	 Storia Cenni storici	leggi
	L'Università' di Padova e' una delle più 'antiche in Europa. Prima di costituirsi in una sede univer permanente, tuttavia, nella citta' medioev	leggi
UNI EN 150 9001:2000	Copyright © 1993-2004 Dipartimento di Ingegneria dell'Informazione <u>Università deoli Studi di Pado</u>	222



IMS Research Group:

- Established in 1987
- Main Research Lines:
 - O Information Retrieval
 - O Digital Library
- Staff:
 - 6 permanent research/teaching
 - O 5 PhD and grant holders
 - O 4 collaborators
- Book on Twenty Years of Research:
 - M. Agosti (Ed). Information Access through Search Engines and Digital Libraries. Springer, Heidelberg, Germany, 2007.

http://ims.dei.unipd.it/





IMS Research Activities and Achievements





- Introduction: University of Padua, DEI, IMS
- Information Access: the Three Main Actors
- Information Access System
 - Centralized Systems
 - O Centralized Systems Distributed Access
 - Distributed Access also Through the Web
- DELOS View of DL, DLS and DLMS
- DELOS Reference Model Overview
- Service Oriented Architecture
- DL Service: the Case of Annotations
- Evaluation of a DLMS
- Personalization through Web Log Analysis









- Introduction: University of Padua, DEI, IMS
- Information Access: the Three Main Actors
- Information Access System
 - Centralized Systems
 - O Centralized Systems Distributed Access
 - Distributed Access also Through the Web
- DELOS View of DL, DLS and DLMS
- DELOS Reference Model Overview
- Service Oriented Architecture
- DL Service: the Case of Annotations
- Evaluation of a DLMS
- Personalization through Web Log Analysis







- Library Automation System or Catalogue Databases System since it manages Bibliographic/Catalogue Databases to support services for:
 - retrieval of catalogue information
 localization of traditional library material
 Management Software



Library Automation System accessible through Internet
 Online Public Access Catalogue (OPAC)

- Digital Library Management System: accessible through Internet and the Web
- Multimedia Contents
- Digital Objects Access through textual data

- Introduction: University of Padua, DEI, IMS
- Information Access: the Three Main Actors
- Information Access System
 - Centralized Systems
 - O Centralized Systems Distributed Access
 - Distributed Access also Through the Web
- DELOS View of DL, DLS and DLMS
- DELOS Reference Model Overview
- Service Oriented Architecture
- DL Service: the Case of Annotations
- Evaluation of a DLMS
- Personalization through Web Log Analysis

DELOS Achievements

- DELOS View Digital Libraries Universe
- DELOS Reference Model
- Service Oriented Architecture
- DelosDLMS
- Evaluation of a DLMS

- Digital Library: A (potentially virtual) organization that comprehensively collects, manages, and preserves for the long term rich digital content and offers to its user communities specialized functionality on that content, of measurable quality, and according to prescribed policies
- Digital Library System: A software system that is based on a (potentially distributed) architecture and provides all the functionalities that are required by a particular Digital Library. Users interact with a Digital Library through the corresponding Digital Library System
- Digital Library Management System (DMLS): A generic software system that provides the appropriate infrastructure to
 - produce a basic Digital Library System that incorporates all functionality that is considered foundational for Digital Libraries
 - integrate additional software offering more refined, specialized, or advanced functionality

DL, DLS, and DLMS: A Three-Tier Framework

Figures extracted from: DELOS Deliverable D1.4.2 – Reference Model for Digital Library Management Systems, September 2006 and The DELOS Manifesto, 2007.

- Introduction: University of Padua, DEI, IMS
- Information Access: the Three Main Actors
- Information Access System
 - Centralized Systems
 - O Centralized Systems Distributed Access
 - Distributed Access also Through the Web
- DELOS View of DL, DLS and DLMS
- DELOS Reference Model Overview
- Service Oriented Architecture
- DL Service: the Case of Annotations
- Evaluation of a DLMS
- Personalization through Web Log Analysis

DELOS Reference Model Overview

- Introduction: University of Padua, DEI, IMS
- Information Access: the Three Main Actors
- Information Access System
 - Centralized Systems
 - O Centralized Systems Distributed Access
 - O Distributed Access also Through the Web
- DELOS View of DL, DLS and DLMS
- The Digital Libraries Universe
- DELOS Reference Model Overview
- Service Oriented Architecture
- DL Service: the Case of Annotations
- Evaluation of a DLMS
- Personalization through Web Log Analysis

□ DL Functionality is available by dedicated services

- DL Applications can be built by combining existing DL services (possibly from different providers), independent of content
- □ Leads to extensible systems that can easily be adapted

Generic DLMS Service

- Have a demonstrator for future Digital Library Management Systems that not only shows new combined text and audiovisual search functionality and personalized browsing by new adaptable information visualization and relevance feedback tools at the interface
- Prove that generic systems can be built that not only support finding of relevant information but also enable to annotate and process found information
- □ Integrate sensor data stream processing
- Allow simple configuration and adaptation while being reliable and scalable

- Introduction: University of Padua, DEI, IMS
- Information Access: the Three Main Actors
- Information Access System
 - Centralized Systems
 - O Centralized Systems Distributed Access
 - O Distributed Access also Through the Web
- DELOS View of DL, DLS and DLMS
- The Digital Libraries Universe
- DELOS Reference Model Overview
- Service Oriented Architecture
- DL Service: the Case of Annotations
- Evaluation of a DLMS
- Personalization through Web Log Analysis

DL Service: the Case of Annotations

- A very well-established practice and widely used
- A way for merging personal contents with the information resources provided by DL
- A means for improving collaborative work practices.

Historical Annotations: Padua

CONSILIARYS - MIGNIFICO-WHER FOTORE MAN PAVLOVAN HVLLEGARDE -DEPRY DX 1 Dies-Bealerthe Derivelling (allan-ALLOINNE AVER OF Q'recurstor bus astrirolenfi aste Wolftange Weie materage of - CONSTLIARIIS Barriele Bellont Tanker Stranger h - PAVIO VAN HVILLEGAR - de Brabanto Bruxel-The lenti, & grass + CHRISTIANO ROSIAN LANCREAUS _ RISCH DAMUE MENA -S. Milenenf: MF -Procuratoribus-Wolffeance Weiß Auth Bastaco & 27.6 a3 Stancredo Zalio Dana si AND NO SHE -SE+1005about here instructions malito intransmorting detil at Chert mon a2 a1

Italia, Padova, Archivio dell'Università di Padova, Archivio antico, *Matricula Nationis Germanicae artistarum*, reg. 465, c. 69v

Historical Glosses: Bologna

Italia, Bologna, Biblioteca Comunale dell'Archiginnasio, Salatiele, *Ars notariae*, ms B 1484, c. 12r

```
Gourd cars gasagest norbad
We cotted the tablinds for education of the
   ing intervention and the set and
 a mouse particul guinom us and 5 run
und so cause north gettin at constitue
nelver quarie cenerary jens net a
mulditte Refur ut agi interior
 cobligation to contrainfitti 5 Squafia
Pant oblato cemalchito cum quis 5
man sun que une forman ap 1.m
ana, to aprolim or patter marby stored
כול בחיפואות השמים שמולים אל
 and net most the childle net were count
ចូលជាពិជាអនុ តាំងស្អោះចាប់ ចាប់អនាំ ប្រទេ
And veneration overal efficience proping
  on of hits from the but frever white
  aut south any auguration and a lot that
 tarb beseeni ut efficien
                  etomologication
are quoted some this filter, a co
      ture and an arrange furthing
11610
CONTRACTOR OWNERS COURT ADDOVIDED
 phone has been undersoome and be
 mentile or ion ut of . openin ci bus
 untermaterie algually form
 is after interior and the state of the parties
riccho. & your furring with Suntr co
  To mins the pursue non batur D
                       or games oches not such a special distant of the sectors y shape.
                               and the faith of the examined a calletter a faith and
                     frifer anbarn mehn mer fant abgemenet gar tit puffern i delige
                      went with the mainhas 1523
```


Annotation Concepts

Document-Annotation Hypertext

Example of Document-Annotation Hypertext

Flexible Annotation Service Tool (FAST)

Main Features

- Flexible and modular architecture deployed as a **component service** of an **Digital Library Management System** (DLMS)
- Full user and group management for improved and finegrained collaboration
- Advanced document/annotation search functionalities based on the hypertext between documents and annotations

Achievements

- FAST as stand-alone prototype
- FAST as Web service into the DelosDLMS

FAST as Stand-alone Prototype

Search by Annotation into DelosDLMS

MCCKII MCCKI
RUN . IUH

DelosDLMS interface has been extended for adding annotations to objects

16RS
U . UNC

Annotations are displayed together with all other meta data of an object

DELOS

NETWORK OF EXCELLENCE ON DIGITAL LIBRARIES

DELOS

Keyword search can be performed on metadata only (marked in gray), on annotations only (marked in green), or on both (marked with blue color)

- Introduction: University of Padua, DEI, IMS
- Information Access: the Three Main Actors
- Information Access System
 - Centralized Systems
 - O Centralized Systems Distributed Access
 - O Distributed Access also Through the Web
- DELOS View of DL, DLS and DLMS
- The Digital Libraries Universe
- DELOS Reference Model Overview
- Service Oriented Architecture
- DL Service: the Case of Annotations
- Evaluation of a DLMS
- Personalization through Web Log Analysis

The evaluation itself of a DLMS turns out to be a scientific activity whose outcomes, such as performance analyses and measurements, constitute a kind of scientific data that need to be properly considered and used for

- the design and
- O development

of DLMS components and services

DIRECT (Distributed Information Retrieval Evaluation Campaign Tool) is a digital library system for managing the scientific data produced during an evaluation campaign

DIRECT has been adopted in
 CLEF 2005
 CLEF 2006
 CLEF 2007

DIRECT Participant Interface

	FOR Cross Languag	ge Evaluation	Fo	rum (CLEF))		particip	ant
			2		Logged as daec	dalus [Logou	at]	
Main						0 N	lessages	
Submission Status	🕝 Track group			🛿 Submission Clo	osing Date	🕙 As:	sessments and Metrics	
	A H-BILLCI FE2006 Ad Vac Bilingual tack			04-jup-20	006		vailable	л,
	ATT DET GEET 2000 - Au-Hot Bringu	ui tusk		04 Juli 20		,	a valiable	
	AH-MONO-CLEF2006 - Ad-Hoc Monoli	ngual task		04-iun-20	06		Vailable	
	AH-MONO-BG-CLEF2006		Ad-Ho	oc Monolingual Bulga track	arian	Download	- 1	
	🖻 Run ID	🗿 Timestamp		💜 Priority	🔯 Topic Fi	elds	🃁 Descriptio	on
	bgFSbg3S View Run Info View Experiment Metrics	viernes 2 de junio at 00H22' CES	de 20 ST	06 2	descrip narrat title	otion :ive e	TDN run	
	bgFSbg2S View Run Info View Experiment Metrics	viernes 2 de junio de 20 at 00H20' CEST		06 ₃	descrip title	otion e	Baseline mandate TD run	
	bgx011bgFW4FS4 View Run Info View Experiment Metrics at 00H24' CEST		de 20 ST	06 1	descrip narrat title	otion tive e	Asymmetr combination o runs	ic ₁f TDN
	AH-MONO-EN-CLEF2006		Ad-H	d-Hoc Monolingual English track Download Track Info as XML				- v

DIRECT Assessor Interface

IKE		for Cross Language F	valuation Forum (CLEF)					
IOEMANDA CHERALIVALIA	N CAMPAGN FOOL						asses	ssor
ORI	NO P.	N LIVORNO-PISA	9:46 21		•• T O	RIN	NO F	D'
			DEC DIAD II			NITI		
	UR. 50	LAUR	REG 9.49 11		• • M I	NIL	JR. 3	20
RAS	CATI		REG 9:52 11	8.	-FR	ASC	AT	
			DID O'FO O		and the second second second			
IUM		JA.	DIR 9.52 2	. 7 .	• F			
				Logge	ed as petya [l	.ogout]		
ssessment	s				9	Messad	26	
Jocobinent						ressug		
			Reload Pools					
cocomonto						0.4-		
Status		📝 Pool Name	🍃 Pool Description			AS 😈	sessed Docu otal Docume	umei ents
0	AH	I-BULGARIAN-CLEF2006	Bulgarian Pool for AH-MONO-BG, a	and AH	-BILI-X2BG	i 17	308 / 17	7 30
	🗑 Topic ID	0	Topic Summary	🔵 Total	\ominus Relevant 🗧	Not Relevant	⊖ Not Assessed	As
	301	Какви продукти на Нестле се продават	по света?	487	2	485	0	
	302	Намерете документи, които описват или дискутират въздействието на бойкоти от страна			6	407	0	
	303	на потребителите. Намерете информация за места, където са издожени италиански картини			9	358	0	
	304	Дайте имената и/или нахождението на местата, които са били посочени от ЮНЕСКО за			14	200	0	
	304	места на световното наследство с изкли	места на световното наследство с изключителна красота или важност.			200	0	
	305	Коя е настоящата цена на нефта?			16	420	0	
	306	Намерете документи, описващи деиността на оаската терористична група ЕТА във Франция.			3	327	0	
	307	Намерете каквато и да е информация за филми, чието действие се развива в Шотландия.			3	271	0	
	308	Намерете документи, които споменават за пълно или частично слънчево затъмнение.			11	472	0	
	309	Какви са опасностите за здравето от нелегалните "твърди" наркотици, като например хероин и кокаин, в сравнение с "леките" наркотици?			25	225	0	
	310	Намерете документи, които описват методи, използвани за обработката или			8	313	0	
	311	Намерете информация и цифри за нивото на безработица в европейските страни.		259	5	254	0	
	312	Какви са последствията от нападенията	а на кучета над хора?	299	26	273	0	
	313	Намерете документи, описващи честван	ия по случай 100-годишнината на някое събитие.	318	42	276	0	
	211	Какви мерки се вземат в Европа, за да	се защитят животните или растенията, които са	400	30	460	0	
	314	изложени на риска да изчезнат?		400	50	100	0	

DIRECT @ TREC

- Introduction: University of Padua, DEI, IMS
- Information Access: the Three Main Actors
- Information Access System
 - Centralized Systems
 - O Centralized Systems Distributed Access
 - O Distributed Access also Through the Web
- DELOS View of DL, DLS and DLMS
- The Digital Libraries Universe
- DELOS Reference Model Overview
- Service Oriented Architecture
- DL Service: the Case of Annotations
- Evaluation of a DLMS
- Personalization through Web Log Analysis

- Analysis of the Web log files in the W3C Extended Log File Format
- Preliminary analysis using both HTTP requests and user surveys gives information on users from different point of view:
 - O Geographical areas
 - O Use of different collections of digital objects
- □ Further analysis is necessary to give:
 - O Better understanding on the usage of digital objects by users
 - Directions towards innovative personalized services

Not only information access but also direct production and contribution by final users

This work was partially supported by the DELOS Network of Excellence on Digital Libraries, as part of the Information Society Technologies (IST) Program of the European Commission (Contract G038-507618)

Thank you for your attention