United Nations
Educational, Scientific and
Cultural Organization

National
Digital Library
of India

INDIAN INSTITUTE OF
TECHNOLOGY KHARAGPUR

UNESCO-NDL INDIA INTERNATIONAL WORKSHOP ON KNOWLEDGE ENGINEERING FOR DIGITAL LIBRARY DESIGN

United Nations
Educational, Scientific and
Cultural Organization

National
Digital Library
of India

INDIAN INSTITUTE OF
TECHNOLOGY KHARAGPUR

UNESCO-NDL INDIA

INTERNATIONAL WORKSHOP ON KNOWLEDGE ENGINEERING FOR

DIGITAL LIBRARY DESIGN

25th - 27th October 2017

New Delhi, India

About the Workshop

The UNESCO-NDL India International Workshop on Knowledge Engineering for Digital Library Design is intended to deliberate on the state-of- the-art technology, practices, and policies as internationally accepted and available in the domain of digital libraries and establish connections among digital library professionals. It is expected that the Workshop will lead to establishment of a forum of professionals of the international digital library community for regular and sustained interactions towards development of the profession.

The participants of the Workshop are comprised of professionals and researchers of the concerned technical and library science domains, senior librarians and digital library professionals of leading libraries.

PATRONS

Kewal Kumar Sharma
Secretary (Dept. of HE), MHRD

sec.dhe@nic.in

R Subrahmanyam
Addl. Secretary (TE)

subrahyd@gmail.com

Partha Pratim Chakrabarti
PI, National Digital Library of India
& Director, IIT Kharagpur

director@iitkgp.ernet.in

Shigeru Aoyagi
Director and UNESCO Representative
to Bhutan, India, Maldives and Sri Lanka

s.aoyagi@unesco.org

Indrajit Banerjee
Director, Knowledge Societies Division,
Communication and Information Sector,
UNESCO

i.banerjee@unesco.org

GENERAL CHAIRS

Al Amin Yusuph
Advisor
(Communication and Information)
for South Asia,
UNESCO New Delhi Cluster Office for
Bangladesh, Bhutan, India, Maldives,
Nepal and Sri Lanka

a.yusuph@unesco.org

Kristian Jensen
Head of Collections & Curation,
British Library, London

Kristian.Jensen@bl.uk

Partha Pratim Das
Jt. PI, National Digital Library of India
& Professor, IIT Kharagpur

ppd@cse.iitkgp.ernet.in

PROGRAM CHAIRS

Jill Cousins
Executive Director,
Europeana Foundation

jill.cousins@europeana.eu

Uma Kanjilal
Director, ICU-TEFED & Professor,
IGNOU, New Delhi

umakanjilal@gmail.com

Plaban Kumar Bhowmick
Co- PI, National Digital Library of India
& Assistant Professor, IIT Kharagpur

plaban@cet.iitkgp.ernet.in

PUBLICATION CHAIRS

Parthasarathi Mukhopadhyay
Associate Professor, Library and Information Science, University of Kalyani, West Bengal
psmukhopadhyay@gmail.com

Ramesh C Gaur
University Librarian, JNU, New Delhi
rcgaur66@gmail.com

ORGANIZING CHAIRS

B. Sutradhar
Co- PI, National Digital Library of India & Librarian, IIT Kharagpur
bsutra@library.iitkgp.ernet.in

P. K. Bhattacharya
Fellow & Head, Library and Information Center, TERI, New Delhi
pkbhatta@teri.res.in

Anirban Sarma
National Programme Officer (Communication and Information), UNESCO New Delhi Cluster Office for Bangladesh, Bhutan, India, Maldives, Nepal and Sri Lanka
a.sarma@unesco.org

PUBLICITY CHAIRS

Shantanu Ganguly
Fellow & Assoc Editor (WDL) TERI
shantanu.ganguly@teri.res.in

Rekha Beri
Documentalist and Public Information Officer, UNESCO New Delhi Cluster Office for Bangladesh, Bhutan, India, Maldives, Nepal and Sri Lanka
r.beri@unesco.org

FINANCE CHAIRS

Nanda Gopal Chattopadhyay
CTO, National Digital Library of India, IIT Kharagpur
nandgc@gmail.com

Ashita Singh
Senior Programme Assistant (Communication and Information), UNESCO New Delhi Cluster Office for Bangladesh, Bhutan, India, Maldives, Nepal and Sri Lanka
a.singh2@unesco.org

Message from Director, Knowledge Societies Division, UNESCO

UNESCO lauds the establishment of the National Digital Library (NDL) for India, supported by the Ministry of Human Resource Development, Government of India and developed by IIT Kharagpur. The NDL is a robust discovery system offering access to a vast range of scholarly research and educational material. A national asset, it provides the basis for numerous future collaborations with institutions and digital libraries worldwide.

The IFLA–UNESCO Manifesto for Digital Libraries, adopted by UNESCO's member states in 2011, notes that: 'The mission of the digital library is to give direct access to information resources ... in a structured and authoritative manner and thus to link information technology, education and culture in contemporary library service.' By integrating digitized and digital contents across educational, research, cultural and heritage institutions, and allowing diverse user groups access to information and knowledge, the NDL goes beyond fulfilling the core mandate of a digital library. It is helping build an inclusive knowledge society, and contributing to the global development goal of ensuring public access to information (SDG 16.10).

Among the most significant threats to digital continuity today is the possibility of losing means of access. UNESCO's Charter on the Preservation of Digital Heritage draws attention to several factors that threaten access, and encourages stakeholder organizations to develop a collaborative, pragmatic and compliant operating environment that will help maximize access to digital heritage. It is to the NDL's credit that within a short span of three years, it has built important linkages with leading institutions and archives both in India and around the world in order to strengthen its technical capacity, expand its content base, and make accessibility more inclusive by promoting access for disadvantaged groups such as persons with disabilities. The NDL's award-winning mobile app is yet another user-friendly platform to access its contents seamlessly.

I am certain that the deliberations at the International Workshop on 'Knowledge Engineering for Digital Library Design' in New Delhi will be most enriching, and that the ideas exchanged will help build enduring knowledge networks. I wish the National Digital Library of India the very best for its future initiatives. all the success.

Indrajit Banerjee

Message from Director and UNESCO Representative for Bhutan, India, Maldives and Sri Lanka

The contribution of digital libraries in stimulating sustainable development and their potential in archiving valuable heritage archives for generations to come cannot be underestimated. Digital libraries such as the National Digital Library of India are already serving as cost-effective, ubiquitous platforms to access knowledge, which support citizens' informed decision-making.

The National Digital Library initiative comes at an appropriate time, when the world is faced with the challenges of providing quality content to a large number of Internet users, the majority of whom are from India and a good number of them from the academic sector. Nearly 90 per cent of the 750 million people who went online for the first time between 2012 and 2015 were from developing economies, with the largest number – at close to 178 million – from India.

Automation in the form of robots, artificial intelligence and the Internet of Things will bring many positive changes to our societies. The new technological developments will also drive many citizens out of many traditional jobs. This will result in the need for massive programmes to retrain citizens for new job profiles. In the current technology-driven and fast-changing world, lifelong learning at all levels is crucial for India and for the world at large to achieve the Sustainable Development Goals. It is for these reasons that the National Digital Library of India occupies a special place in skilling the 1.3 billion citizens of India.

The UNESCO constitution commits our organization to advance mutual knowledge and understanding of peoples of the world, and "promote the free flow of ideas by word and image". UNESCO also recognizes the importance of promoting multilingualism and equitable access to information and knowledge, especially in the public domain. The National Digital Library is built on the same shared values by providing multilingual content and open access to information and knowledge.

I wish to assure National Digital Library and the Ministry of Human resource Development of our continued collaboration and commitment for this noble initiative, which is in line with our organization's mandate. I also hope that UNESCO's contribution to this workshop will lead to the intended sharing of international best practices, policies and state-of-the-art technology in the domain of digital libraries.

Shigeru Aoyagi

भारतीय प्रौद्योगिकी संस्थान खड़गपुर
खड़गपुर-७२१ ३०२
Indian Institute of Technology Kharagpur
Kharagpur – 721302

प्रोफेसर पार्थ पी. चक्रवर्ती, एफएनए, एफएएससी, एफएनएई
निदेशक
Prof. Partha P. Chakrabarti, FNA, FASc, FNAE
J. C. Bose Fellow
Director
Professor, Department of Computer Science & Engineering

Message from Patron

India is a country of enormous diversity. For ages, she has observed harmonious co-existence of different cultures, ethnic groups and others. Consequently, great variation in creative realm and intellectual property is prevalent amidst this diversity. With rapid adoption of digitization and ICT, access to knowledge has become easier pervading space and time. However, dissemination of knowledge has been performed in distributed manner so far. The call of the hour is to leverage ICT to string together scattered intellectual property and creative works by respecting the autonomy of the respective knowledge sources. This is where digital library technology plays the most crucial role by permeating spatial, cultural and socio-economic barriers.

Knowledge engineering lays the foundation of any intelligent design. Digital library technology is no exception. Over the years, digital library fraternity has come up with standards and recommendations to represent and organize digital resources. While the legacy systems have widely adopted these standards, new digital libraries are embracing advances in semantic web and linked data technologies.

While knowledge engineering helps in modelling the digital resources, search or retrieval systems focuses on satisfying information needs of the users by consulting structured knowledge about resources. Though retrieval systems are in place for decades, sparseness in metadata specification poses great challenge to traditional information retrieval algorithms. Diversity in language demands the retrieval systems to deal with multi-lingualism and cross-lingualism.

Satisfying information needs of users with varying background and intent is another important challenge in digital library design. Keeping in mind the demographic and temporal variations in information needs, modern digital libraries are adopting user-centered design paradigm. Research outcomes in social science and psychology domains should be integrated to come up with strategies to engage users. On the other hand, these strategies should guide the design of effective user interfaces to access resources in digital library. The design should explore alternative interfaces to cater to the differently-abled population.

The access policies in digital library design should stress upon democratization of knowledge to break free from the divide that exists in form of demography and socio-economic background. It will be interesting to see how do digital libraries ride the wave of open access by adopting evolving space of open licensing policies like creative commons, OERs and others. It is also worth exploring the implications of open access policies in digital library technology space.

That UNESCO-NDL International Workshop on Knowledge Engineering for Digital Library Technology has rightly taken a step forward by hosting leading experts in digital library and related domains to deliberate on the important challenges and the way forward. I am confident that the library science and digital library fraternity will be immensely benefited by the compact program that is comprised of several keynotes and panels. I wish you all engrossing and productive deliberations. I wish the workshop all the success.

Partha P. Chakrabarti

Office : Phone : +91 3222 282002 **Fax :** +91 3222 282000 **e-mail :** director@iitkgp.ernet.in; ppchak@cse.iitkgp.ernet.in
Residence : Director's Bungalow, IIT Campus, Kharagpur-721 302, **Phone :** +91 3222 282003 / 277201

Message from Honorable Director, IIT Kharagpur

Partha Pratim Chakrabarti

Welcome from General Chairs

On behalf of the Organizers of the UNESCO-NDL India International Workshop on Knowledge Engineering for Digital Library Design we heartily welcome you to the event.

UNESCO is committed to unrestricted pursuit of knowledge for the people of the world through educational, scientific and cultural means and accordingly is championing open access in learning parlance. Digital technology is bringing a paradigm shift in the way mankind acquires knowledge and thus digital libraries are becoming a very effective avenue to realize universal open access of the learning contents. It is in this context UNESCO New Delhi Office has collaborated with the National Digital Library of India to actively promote the cause. The International Workshop on Knowledge Engineering for Digital Library Design is being organized to take this movement forward. The Workshop is being held during International Open Access Week during the week of 23rd October 2017 thus bringing further relevance to the event.

The National Digital Library of India (NDL India aka NDLI) (https://ndl.iitkgp.ac.in) was initiated in April 2015 as a Pilot project by the Ministry of Human Resource Development (MHRD), Government of India, under its National Mission on Education through Information and Communication Technology (NME-ICT) Mission to build the framework of a digital library that can be scaled up with respect to content volume and diversity, over time, to a full-fledged national digital library. The project is being executed by Indian Institute of Technology Kharagpur (IIT Kharagpur). The pilot went live about a year back and currently hosts more than 12 million contents in more than 100 languages, sourced from about 150 institutes/ publishers. It is also available as a Mobile App (Android and iOS). More than 2 million users are currently registered in the library. The motto of NDLI is "Inclusive" and "Open". India is multi-lingual nation and it is in this perspective User Interface of NDLI has been planned for all major Indian languages.

The Workshop has been conceived to deliberate on the state-of-the-art technology, practices, and policies as internationally accepted and available for digital library design. It is going to have 22 speeches spread over 9 sessions and 3 panel discussions. There are about 30 speakers and panelists from a dozen of countries and about 200 participants are going to attend the Workshop. The themes of the Workshop have been carefully chosen to address the knowledge needs of the digital library designer community and cover Digital Library Technologies, Metadata Engineering, Linked Open Data, Academic Search & Cross-lingual Retrieval, Personalization of Search Results, Designing for the User, User Engagement, Open Access & Licensing, and Rights Standardization & Document Preservation.

We expect the participants will find the event a good interactive platform for furtherance of knowledge for digital library design and eventually will establish a forum of professionals of the international digital library community for regular and sustained interaction towards development of open knowledge and create a new paradigm in education.

Al Amin Yusuph

Partha Pratim Das

Kristian Jensen

Message from Program Chairs

It is a great pleasure to welcome you all in UNESCO-NDL India International Workshop on Knowledge Engineering for Digital Library Design. Last few decades have observed a phenomenal growth in digital library technologies built over the foundations like information retrieval, knowledge management, metadata engineering, human computer interaction and likes. Several digital libraries around the world have emerged focusing on collections from different domains like cultural heritage, education etc. Though there are certain commonalities in design goals and approaches, there are interesting differences as well. This workshop aims at bringing together professionals from digital library technology and related domains to deliberate on foundational standards, technologies and to provide glimpses of future ahead.

The workshop is organized around a number of keynote talks and panels categorized into four broad themes. The 'Digital Library Technologies' theme focuses on core technologies involved in design of different digital libraries around the world and how the design goals can be better achieved using modern technologies. This theme will host keynote speeches by technology experts from digital libraries DigitalNZ, Tainacan, Trove, Deutsche Digitale Bibliothek and Europeana. Standardization for knowledge organization has been one of the primary concerns in digital library domain as it provides way for knowledge sharing and interoperability. The 'Standards and Interoperability' theme focuses on core aspects metadata reuse, linked data principle and their relevance towards networking in digital library. The 'Search & Information Retrieval' theme deliberates on challenges and technologies in regard to academic or scholarly search systems like Microsoft Academic and Google Scholar. Talks in this theme also focus on multi-lingual and cross-lingual search systems that are integral parts of digital libraries hosting multi-lingual resources. Emerging search and discovery systems have interesting implications on Academic Search Engine – Discovery Tool – Publisher ecosystem. Modern software systems place users at the centre of design. Currently, digital libraries are widely adopting this paradigm with two broad design goals: Drawing user attention, Sustained user engagement. While deploying appropriate strategies like crowdsourcing, collaboration help in drawing user attention, rich user experience design ensures sustained user engagement. 'Designing for the User' theme will host talks on content strategy and user experience design for different user categories. Rights and access policy of digital resources has become an important issue in content acquisition and dissemination strategy in digital library. Democratization of knowledge through open access policy is gaining enormous support. The 'Access & Rights' theme focuses on future of open access culture and its implication in digital library design. This theme also deliberates on standardization of rights statements so that they are interpreted unambiguously across systems. The workshop will be concluded by a panel discussion on collaboration and networking among existing digital libraries towards effective knowledge organization, sharing and dissemination.

The members of the program committee owe sincere thanks to the esteemed speakers, panelists and session chairs for contributing to the objective of the worshop. We look forward to an exciting workshop with plethora of insightful deliberations, engaging interactions and sharing of knowledge. Finally, we hope that you will enjoy your visit to New Delhi.

Jill Cousins

Uma Kanjilal

Plaban K Bhowmick

PROGRAM SCHEDULE

WEDNESDAY: 25TH OCTOBER

09:00-10:00AM	REGISTRATION	
10:00-10:30AM	WELCOME ADDRESS	
10:30-11:00AM	TEA BREAK	
11:00-11:40AM	Digital Initiatives and NDL of India	Partha Pratim Das
11:40-11:55AM	DigitalNZ 'Supplejack': Open-Source, Metadata Aggregation, Discovery and Management Software	James Robertson
11:55-12:10PM	TAINACAN Project - Experiences and results using WordPress to develop a Digital Repository	Dalton Martins
12:10-12:25PM	Digital Library Technologies: TROVE	Brendon McKinley
12:25-12:40PM	Upgrading Digital Library Technology to the Age of Analytics – The Case of the German Digital Library	Stephan Bartholmei
12:40-01:00PM	Technology is Not Enough - Reflecting about Digital Library Technologies	Pavel Kats
01:00-02:00PM	LUNCH	
02:00-02:45PM	Metadata Reuse: A Discovery Framework for Authority Recommendation and ILS-DI based OPAC Functionalities	Partha Sarathi Mukhopadhyay
03:00-04:00PM	INAUGURATION	
04:00-04:20PM	TEA BREAK	
04:20-05:20PM	PANEL 1: If I were to do it all again what would I use	Dalton Martins (TAINACAN), James Robertson (DigitalNZ), Stephan Bartholmei (DDB), Pavel Kats (Europeana), Brendon McKinley (Trove), Moderator: Partha Pratim Das
05:20-06:00PM	Linked Data – Where, What, Why, What	Richard Wallis
06:00-06:40PM	A Distributed Network of Heritage Information	Enno Meijers

THURSDAY: 26TH OCTOBER

09:00-09:40AM	Handling Noise in Scholarly Search: Challenges & A Case Study	Anurag Acharya
09:40-10:20AM	Academic Search – Challenges and Possibilities	Mihaela Vorvoreanu
10:20-10:40AM	TEA BREAK	
10:40-11:30AM	Multilingual and Crosslingual Retrieval in Digital Libraries	Sudeshna Sarkar
11:30-01:00PM	PANEL 2: Academic Search Engines, Discovery Tools and Publishers - Collaboration or Competition	Mihaela Vorvoreanu, Paul D Clough, Paul Nieuwenhuysen, John Akeroyd, Balaji Devarajan Jayaraman Moderator: Jagdish Arora
01:00-02:00PM	LUNCH	
02:00-02:40PM	Making Users Stakeholders – Strategies to Engage and Collaborate	Jill Cousins
02:40-03:20PM	User-informed Design for Search and Discovery in Digital Cultural Heritage	Paul D Clough
03:20-04:00PM	Digital Library and Accessibility	Anupam Basu
04:00-04:20PM	TEA BREAK	
04:20-05:00PM	OCR for Indian Languages and Beyond	C V Jawahar
05:00-05:40PM	Crowdsourcing Based Curation and User Engagement in Digital Library Design	Rose Holley
05:40-06:20PM	Online Learning and Open Educational Resources - Challenge of the Future (with a glimpse at the past)	Rory McGreal
06:20-07:00PM	HIGH TEA	
07:00-08:00PM	CULTURAL PROGRAM	
08:00-09:00PM	BANQUET DINNER	

FRIDAY: 27TH OCTOBER

09:30-10:10AM	Rethinking the future of Openness	Lawrence Liang
10:10-10:50AM	Delocalising Knowledge..... IPR Issues hold the Keys	Prabuddha Ganguli
10:50-11:10AM	TEA BREAK	
11:10-11:50AM	Do the Right thing – Interoperable Rights Statements	Jill Cousins
11:50-01:00PM	Preservation of documentary heritage: UNESCO Experience	Indrajit Banerjee
01:00-02:00PM	LUNCH	
02:00-04:00PM	PANEL 3: Networking and Collaboration among DLs of the World	Europeana (Jill Cousins), NDLI (Partha Pratim Das), DigitalNZ (James Robertson) TAINACAN (Dalton Martins), South Africa (Douwe Drijfhout), Nepal (Upendra Prasad Mainali), Bangladesh (Mezbah-ul-Islam), Moderator: Ramesh C Gaur
04:00-04:20PM	Vote of Thanks	Partha Pratim Das

Theme:

DIGITAL LIBRARY TECHNOLOGIES

Talk 1: Digital Initiatives of NDL of India

Talk 2: DigitalNZ 'Supplejack': Open-source, Metadata aggregation, Discovery and Management Software

Talk 3: TAINACAN Project - Experiences and results using Wordpress to develop a Digital Repository System

Talk 4: Digital Library Technologies - TROVE

Talk 5: Upgrading Digital Library Technology to the Age of Analytics – the Case of the German Digital Library

Talk 6: Technology is not enough - Reflecting about Digital Library Technologies

PANEL 1: IF I WERE TO DO IT ALL AGAIN WHAT WOULD I USE

DIGITAL INITIATIVES AND NDL OF INDIA

PROF. PARTHA PRATIM DAS

Jt-PI, National Digital Library of India
Professor, IIT Kharagpur

25TH OCT, 11:00-11:40AM

Dr. Partha Pratim Das is the Joint Principal Investigator of National Digital Library of India project of MHRD, Govt. of India and leads the initiative to integrate the Digital Repositories of various Institutions, agencies and publishers across India. He is a Professor at the Department of Computer Science and Engineering, IIT Kharagpur, heads the Rajendra Mishra School of Engineering Entrepreneurship and is the Professor-in-Charge of the upcoming Research Park of IIT Kharagpur at Rajarhat, Kolkata.

Dr. Das received his BTech, MTech and PhD degrees in 1984, 1985 and 1988 respectively from IIT Kharagpur. He served as a faculty in Department of Computer Science and Engineering, IIT Kharagpur from 1988 to 1998. In 1998, he moved to the industry and served in senior management / director positions till 2011. His current interests include Human-Computer Interaction, Computer Analysis of Indian Classical Dance, Technology-Enabled Education, and Software Engineering.

Dr. Das has received several recognitions including UNESCO/ROSTSCA Young Scientist (1989), INSA Young Scientist Award (1990), Young Associate-ship of Indian Academy of Sciences (1992), UGC Young Teachers' Career Award (1993), INAE Young Engineer Award (1996), Interra Special (Process) Recognition (2009), and Interra 10 Years' Tenure Plaque (2011). He served the International Conference on VLSI Design & Embedded Systems in as General Co-Chair in 2005 and as Program Co-Chair in 2016. He is currently the Editor-in-Chief of The Journal of Institution of Engineers: Series B, reviewer for Pattern Recognition Letters and a Review Writer for ACM Computing Reviews. He is a member of Association of Computing Machinery (ACM), The Institute of Electrical & Electronics Engineers (IEEE), and Indian Unit for Pattern Recognition and Artificial Intelligence (IUPRAI).

ABSTRACT

Advances in Digital Technology, specifically in the field of Knowledge Engineering powered by techniques of Artificial Intelligence and Machine Learning, have been changing the face of education and learning across the world since the turn of the millennium. Ministry of Human Resources Development, Govt. of India, realizing the opportunities and the potential, launched its National Mission on Education through Information and Communication Technology (NMEICT) in 2009 to support Technology Enhanced Learning (TEL) schemes to initiate a paradigm shift in the education landscape in India. Under NMEICT several projects like NPTEL (National Programme on TEL), NPTEL-NOC, Virtual Laboratory, e-Yantra, etc. have been taken up for nearly a decade. Through these various high quality contents have been created, variety of distance mode trainings have been imparted, and infrastructure created at wide range of institutions. Yet, the goal of taking education to every citizen, to every corner of the country remains elusive.

Higher Education in India: Vision 2030 Report from **FICCI Higher Education Summit 2013** observes that *"Despite many new national missions/programs and reforms agenda, by both the central and state governments with private sector intervention, the higher education sector is in a state of complete flux".* Further, it notes – *"The Indian higher education system has undergone massive expansion to become the largest in the world enrolling over 70 million students. Such expansion would have been unimaginable without the extensive use of ICT tools. To illustrate, if India were to create this additional capacity through increase in brick and mortar institutions alone, it would have had to build six universities and 270 colleges each and every month in the last 20 years (circa 2030) – a feat that would have been impossible to achieve with India's limited resources. Instead, India chose to go the MOOCs way."* In this background MHRD decided to strengthen its pan-India Digital Initiatives in Education and National Digital Library in India (NDLI) was conceptualized and initiated in 2015 to create a National Knowledge Asset as the key driving force for education, research, innovation, and knowledge economy in India. Formalization and strengthening of NPTEL-NOC – India's own MOOCs – was launched as SWAYAM (Study Webs of Active Learning for Young Aspiring Minds) in July 2017. To widen and strengthen the reach of TEL, 32 Direct-To-Home (DTH) TV Channels were planned as SWAYAM Prabha. NDLI has come live on web and mobile earlier this year and by December 2017 SWAYAM and SWAYAM Prabha are going to be fully integrated with every university in the Country. MHRD has already mandated the inclusion of 20 MOOCs credits in university programs. With NDLI acting as a virtual knowledge container with wealth of learning resources for the students and learners across geographical and cultural boundaries education in India is on the cusp of a paradigm shift. This talk will quickly walk the audience through some of the Digital Initiatives and elucidate the role of NDLI in this paradigm shift.

"With NDLI acting as a virtual knowledge container with wealth of learning resources for the students and learners across geographical and cultural boundaries education in India is on the cusp of a paradigm shift."

JAMES ROBERTSON

Systems Manager for DigitalNZ, National Library of New Zealand

25TH OCT, 11:40-11:55AM

James Robertson is the Systems Manager for DigitalNZ at the National Library of New Zealand. He has extensive experience in the Agile management of Digital products, services and people. For the last 10 years he has held various Software Development Management (SDM) positions including Bank of New Zealand (BNZ), Open Polytechnic and web consultancy firm 3months.com.

His current responsibilities at the National Library include the management and ongoing development of award-winning, open-source software "Supplejack", the superb metadata aggregation, management and discovery tool that underpins the DigitalNZ API and web services, as well as a growing number of similar services around the world.

In the past he has had varied roles as a programmer, developing web-applications and engineering software on embedded systems, and has worked for small and medium enterprises (SMEs) and government in NZ, as well as large corporates in the UK and Switzerland, in both permanent and contract roles.

DIGITALNZ 'SUPPLEJACK': OPEN-SOURCE, METADATA AGGREGATION, DISCOVERY AND MANAGEMENT SOFTWARE

ABSTRACT

Launched by the National Library of New Zealand in 2008, DigitalNZ helps people find, share, and use NZ digital content through it's website (digitalnz.org) and public API (api.digitalnz.org). Our core service is a search infrastructure that aggregates material from over 200 content partners across NZ and around the world. We connect kiwis with over 30 million photos, audio and video files, articles, and documents, serving over 10 million queries a month.

As well as our public websites, DigitalNZ provides open data services via our public API. The DigitalNZ ecosystem is one of the NZ government's most sophisticated online services; acquiring and enhancing metadata from more than 200 data sources on a 24/7 basis, then sharing that data via API to power websites and applications around NZ.

'Supplejack' is DigitalNZ's custom-developed, open-source software for collecting, enhancing, searching, sharing and managing metadata. It can harvest from many disparate sources using a variety of formats including: OAI, Sitemaps, RSS, XML and HTML. It is built using high-performance, open-source tools such as: Ruby on Rails (application language and framework), MongoDB (scalable, metadata document database), Solr (fast, enterprise search platform) and Redis (in-memory data structure store).

Supplejack is a coherent set of Ruby on Rails applications that make life easy for harvest operators. Its key strengths are:

- Flexibility and power of it's scriptable, harvesting 'parsers', utilising both standard Ruby syntax as well as our simplified Domain Specific Language (DSL)
- Best of breed open-source tools enabling data storage at scale and high performance search indexing and retrieval
- Simple but rich Restful API enabling data to be shared and interacted with across a wide range of sites and services

The sophisticated Supplejack 'Manager' dashboard allows non-technical users to:

- Create, edit and test new data sources and parser scripts, including 'live preview' of metadata mappings and enrichment changes
- Keep track of scheduled harvests on both staging and production environments
- Interrogate and moderate individual records

Supple jack is now in use by a variety of organisations around the world and is under constant development. James Robertson, DigitalNZ Systems Manager, will be talking more about the technology behind Supplejack and DigitalNZ as well as plans for the future.

Available at http://digitalnz.github.io/supplejack/

PROF. DALTON MARTINS

Professor, Faculty of Information & Communication,
Federal Univeristy of Goiás

25TH OCT, 11:55AM-12:10PM

Professor in the Information Management course and Postgraduate Program in Communication PPGCOM (Master degree) of the Faculty of Information and Communication of the Federal University of Goiás. He holds a degree in Electrical Engineering from the State University of Campinas (2002) and a Masters in Engineering Computing by the State University of Campinas (2004). PhD in Information Sciences from University of São Paulo (2009-2012), working on the theme of mapping, structural analysis and dynamics of Social Networks in distributed digital environments. Has experiences in the areas of inclusion and digital culture, analysis of social networks, metric studies and human dynamics. He has been working in research at the interface of the Communication and Information areas, focusing on factors and conditions that favor the formation of intelligent collectives and researches with data science applications (machine learning and data mining) in problems involving public policies, media and social participation.

TAINACAN PROJECT: EXPERIENCES AND RESULTS USING WORDPRESS TO DEVELOP A DIGITAL REPOSITORY SOFTWARE

ABSTRACT

The main objectives of this talk are present Tainacan project, a Brazilian experience in developing a free software based on WordPress, to support a public policy from Ministry of Culture in:

- Develop a user-friendly, customizable and easy to manage platform, Tainacan, aimed at managers and users of digital collections;
- Implementing social participation dynamics and stimulating collective intelligence in the management of collections, allowing new ways of inclusion of users and new modes of participatory management;
- To serve as the basis for the implementation of a network service, where cultural institutions and projects can access Tainacan without the need of installation, creating their account, making their collections available and managing in a decentralized manner;
- Facilitate the integration of the different existing collections in Brazil and generate a single search field to facilitate access to the various contents already digitized and available in the country;
- Facilitate and promote the management of museum assets in its version Tainacan + Museums.

> “The TAINACAN Project is a Brazilian experience in developing a free software based on Wordpress, to support a public policy from Ministry of Culture.”

BRENDON McKINLEY
Director, Products Branch,
National Library of Australia, TROVE
25TH OCT, 12:10-12:25PM

DIGITAL LIBRARY TECHNOLOGIES:TROVE

Trove helps you find and use resources relating to Australia. It's more than a search engine. Trove brings together content from libraries, museums, archives, repositories and other research and collecting organisations big and small. A list of contributors and partners whose collections are included in Trove can be viewed online.

Trove's origins can be traced back to a project launched by the National Library of Australia in August 2008. Its aim was to build a portal for all of the Library's online discovery services, including the Register of Australian Archives and Manuscripts, Picture Australia, Libraries Australia, Music Australia, Australia Dancing, PANDORA web archive, ARROW Discovery Service and the Australian Newspapers Beta service.

Today Trove is transformed, growing far beyond its original purpose and becoming many things to many people: a community, a set of services, an aggregation of metadata, and a growing repository of full text digital resources. Trove is a platform on which new knowledge is being built. It is a collaboration between the National Library, Australia's State and Territory libraries and hundreds of cultural and research institutions around Australia, working together to create a legacy of Australia's knowledge for now and into the future.

Source: http://trove.nla.gov.au/general/about

“Today Trove is transformed, growing far beyond its original purpose and becoming many things to many people:
a community, a set of services, an aggregation of metadata, and a growing repository of full text digital resources.”

STEPHAN BARTHOLMEI

Head of Product Development & Innovation, German Digital Library

25TH OCT, 12:25-12:40PM

Stephan Bartholmei is a physicist turned software developer turned product owner. He has worked in academic research and education, in commercial software development and intellectual property rights. In 2013 he has joined the German National Library where he is the head of product development and innovation for the German Digital Library. Stephan is the co-founder of Germany's cultural data hackathon codingdavinci.de and is a Members Councillor for the Europeana Network Association.

ABSTRACT

The first wave of cross-domain digital libraries like Gallica, Trove, Europeana or the German Digital Library (in German: "Deutsche Digitale Bibliothek", short: "DDB") have extended the 1980s OPAC concept to other cultural heritage sectors. Internally, their systems usually comprise a metadata repository and media file store, a search engine set up on an inverted index and an ETL pipeline to feed the repository and search index. Often, their architecture was modeled after the ISO OAIS reference model.

UPGRADING DIGITAL LIBRARY TECHNOLOGY TO THE AGE OF ANALYTICS: THE CASE OF THE GERMAN DIGITAL LIBRARY

After a decade the gained experience shows that a metadata catalogue isn't what the users needed nor wanted. Truly successfull products like Trove's newspapers portal indicate that they are primarily interested in content.

The advent of multisided platforms for retail products (Amazon), social exchange (Facebook, Twitter, Pinterest, ...) and creators (Youtube) demonstrated the importance of precisely knowing your users needs. Accordingly the business models of these platforms focus above all on collecting and analyzing usage data.

A couple of years ago this insight has trickled down into the strategic discussions at the Digital Libraries, whose architectures weren't yet equipped to handle this kind of data.

In my talk I will present the conclusions the German Digital Library drew for her 2020 strategy and how they have influenced our ongoing architecture revamp. When we decided to switch to a state-of-the art NoSQL data store and distributed processing engine our first concern were speed improvements for the data ingestion processes, which have had become slower and slower with the growth of our repository and added features in our portal www.ddb.de, and ensuring the operability up to 100 million or more digital objects.

Apart from performance and scalability concerns our new architecture will enable us to serve new and exciting use cases like semantic enrichments, full-resolution media file storage, metadata and media file analysis with machine learning methods, data warehousing with data inputs from lots of different sources, most importantly usage data from our tracking systems, but also from data hubs we connect to-like the upcoming German School Cloud

We expect data analysis to become a crucial success factor for Digital libraries. Apart from the aforementioned usage statistics analysis, it will enable us to give feedback on metadata quality to our data providers, generate and enrich metadata by feature detection in media assets and full texts and thus improve the information retrieval and search experience for our users. Similarly, the presentation of our aggregated data with state-of-the-art methods like interactive data visualizations and other "generous" interfaces will require considerable computing power. We need to start building the necessary systems and knowledge in our organisations right now.

OPAC = Online Public Access Catalogue
ETL = Extract, Transform and Load
OAIS = Open Archival Information System

PAVEL KATS

Chief Technical Officer, Europeana Foundation

25TH OCT, 12:40-1:00PM

As Europeana´s CTO, Pavel oversees technology related activities in Europeana, which includes infrastructure planning, architecture design, software development, and innovation. Following Europeana´s motto to transform lives with culture, Pavel is interested in the potential of contemporary information technologies applied in the cultural domain to change our society for better. Pavel holds a masters degree in Computer Science from the Hebrew University of Jerusalem.

ABSTRACT

Digital libraries have come a long way in the last decade. From natural but often sporadic extensions of digital systems traditionally employed in libraries, museums, and archives -- primarily, content management systems (CMS) -- they grew into complex and heterogeneous software suites with a diverse set of capabilities designed to serve the continually changing needs of external and internal users.

TECHNOLOGY IS NOT ENOUGH: REFLECTING ABOUT DIGITAL LIBRARY TECHNOLOGIES

During this decade, Europeana has had a unique vantage point to observe the development of digital libraries: we have worked directly and indirectly with many of them to aggregate data; partnered with others to exchange ideas and experience; and -- last but not least -- developed a version of one for our own needs. In this session, I am going to draw on our multi-year experience to offer a holistic outlook on the technological landscape of digital libraries, analyse its evolution in the recent years, discuss how it affected our technological choices, and offer some reflections on the future of the domain.

My starting point will be the technological portfolio that Europeana has developed during these years and the factors which shaped its development. By and large, this portfolio consists of an ingestion toolkit, core storage, data model, search, digital media storage, APIs and the access layer, and presentation.

While these and related components are often discussed in isolation in the context of digital libraries, I'll argue that it is their interdependence and interaction in the context of a wider business strategy, that determines their effectiveness.

As digital libraries evolve, they are increasingly expected to be preoccupied with finding answers to questions shaping the design and development of key pieces of their technological puzzle such as:

How does the data acquisition approach, put to practice using ingestion tools, serve the data quality- and usability-dependent aspects of the overall strategy? What is the right balance between the breadth of data coverage and its semantic interoperability? How does the inclusivity approach to interoperability impact the efficiency of search, a staple of today's user experience? What kind of requirements do state-of-the-art user engagement tools, such as digital exhibitions or semantic-driven navigation and discovery, present to the data and access layers of the system? What is the role of LOD in all this?

My position will be that answers to such questions should be driven by the strategy rather than technology function in digital libraries. As digital libraries come in different shapes and flavours -- to a large extent determined by their funding and governance setup -- they should identify the preferred configuration of several parameters determining their brand and strategy prior to making technological choices. I'll suggest that we, as a technological community, should do more to clearly describe these configurations and what they imply and leverage the considerable body of experience accumulated in recent years to support the decision making process in the digital libraries.

Panel 1:

If I were to do it all again what would I use?

25th October
4:20PM - 5:20PM

The Panelists

PROF. DALTON MARTINS
TAINACAN

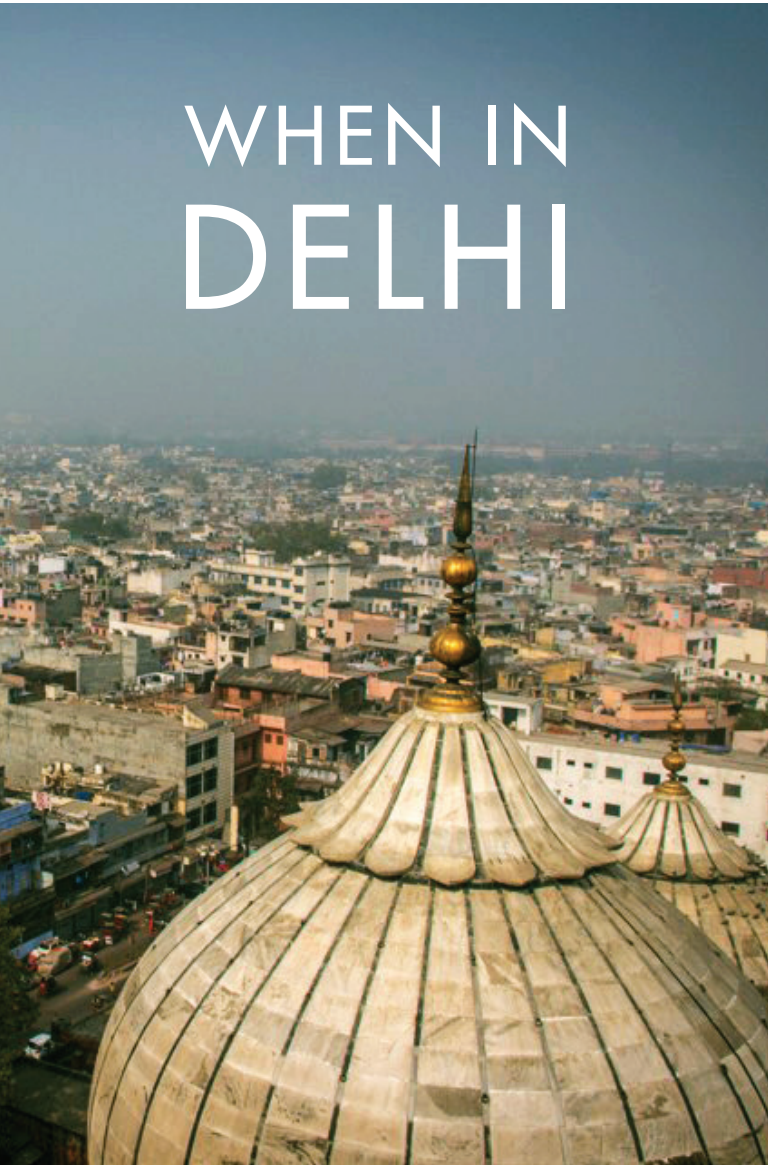
JAMES ROBERTSON
DigitalNZ

STEPHAN BARTHOLMEI
Deutsche Digitale Bibliothek

PAVEL KATS
Europeana Foundation

BRENDON McKINLEY
TROVE

MODERATOR:
PROF. PARTHA PRATIM DAS
National Digital Library of India



PLACES WORTH THE VIEW:

Tughlaqabad Fort

Built in 1321 by Ghyas-ud-din Tughlaq, it is considered to be the third historic city of Delhi. Abandoned later in 1327, legend has it that its construction was cursed by the great mystic and Sufi saint Hazrat Nizamuddin Auliya. Today it lies, a ruined fort, derelict but atmospheric and is a beautiful excursion from the heart of city.

India Gate

India Gate houses the Amar Jawan Jyoti under a marble archway and serves as India's Tomb of the Unknown Solder. This is a favourite spot of the citizens of Delhi for an evening drive with its extensive avenues and indulging in ice-cream sundaes.

Red Fort

Built in 1639 by the Mughal Emperor Shah Jahan, The Red Fort is so named because of its massive enclosing walls of red sandstone and served as the main residence of the Mughal emperors for over 200 years. Every year on 15th August, India's Independence Day, celebrations are held at this fort.

ABOUT ART & CULTURE:

The National Museum and the NGMA

The National Museum and the National Gallery of Modern Art (NGMA), amongst themselves chronicles the artistic and cultural legacies of India. With over 200,000 works of art, the National Museum has on display a span of over 5000 years of history. The NGMA has in its collection over 14,000 artworks including paintings and sculptures by both foreign and Indian artists who have shaped the contemporary art ofthe nation.

TO EAT AND SHOP:

Chandni Chowk

Chandni Chowk and the narrow lanes near Jama Masjid houses some of the finest eateries in the city. A melting pot of people and cultures over centuries, it is at the heart of old Delhi and serves everything from kebabs to parathas (fried flatbread) filled with cream.

Connaught Place

Conceived as the showpiece of Lutyens' Delhi, CP as it is fondly known, is a place where the city comes to shop, eat, drink and have a stroll through the Georgian colonnades and concentric layers of shopping arcades. It houses the iconic Wenger's, the oldest bakery in New Delhi and one of the first shops to open in Connaught Place.

Hauz Khas Village

With the historic Hauz Khas Complex at its heart, this is a neighbourhood in South Delhi known for its cosmopolitan vibe and numerous restaurants and bars dotting the narrow lanes. Modernity sits cheek to jowl with the medieval complex consisting of a water tank, seminary, mosque and pavilions tracing its roots to the 13th century.

Theme:

STANDARDS & INTEROPERABILITY

Talk 1: Metadata Reuse: A Discovery Framework for Authority Recommendation and ILS-DI based OPAC Functionalities

Talk 2: Linked Data – Where, What, Why, What

Talk 3: A Distributed Network of Heritage Information

METADATA REUSE: A DISCOVERY FRAMEWORK FOR AUTHORITY RECOMMENDATION AND ILS-DI BASED OPAC FUNCTIONALITIES

DR. PARTHASARATHI MUKHOPADHYAY

Associate Professor, Dept. of Library & Information Science, University of Kalyani

25TH OCT, 2:00-2:45PM

Parthasarathi Mukhopadhyay is presently working as Associate Professor in Library and Information Science under University of Kalyani. The research domain of Dr. Mukhopadhyay includes application of open source software and open standards in designing library services. He has contributed a total of five books, 60 research papers and many courseware developed for IGNOU. He was a member of UNESCO curricula development committee on Open Access for research scholars and library professionals and developed one exclusive module on Open Access. He is presently running two courses on SWAYAM platform of MHRD, Govt. of India.

ABSTRACT

Metadata holds the key of success in our digital endeavor if supported with suitable interoperability standards. The integration of domain-specific metadata schemas and interoperability standards in a discovery system may lead to an array of user-centric information services required for a pan India project like NDL. This presentation proposal explores the possibility of the application of ILS-DI specifications in designing a prototype union catalogue as discovery system that can provide real-time item level status along with other OPAC functionalities like placement of hold, selection of preferred date range for holds, choice of pickup location etc. This prototype discovery framework is essentially based on open source software and open standards. It applies Koha ILS (different instances) as data provider layers, OAI/PMH as protocol for data harvesting, VuFind as discovery layer and ILS-DI based multibackend driver for communicating different ILS instances (representation of different libraries) to implement the desired OPAC functionalities in union catalogue.

The unique feature of the framework is that it can be applied to include any OAI/PMH compatible ILS and any discovery system. The real magic of the framework is the ILS-DI based multibackend driver, which reuses metadata on the top of OAI/PMH derived datasets for real-time content negotiation. Another interesting attribute of this prototype discovery framework is the integration bibliographic data and authority data (name authority and subject authority formatted in MARC 21 authority data standard) in real-time to generate a recommendation system in end user search interface to support efficient retrieval of relevant resources.

RICHARD WALLIS
Technical Consultant,
Independent Professional, UK

25TH OCT, 5:20-6:00PM

LINKED DATA: WHERE, WHAT, WHY, WHAT

Richard established Data Liberate having been with the UK's leading Linked Data and Semantic Web technology company, Talis, for over twenty years. This coupled with his passion for and engagement with new and emerging technology trends, gives him a unique perspective of the issues challenging information professionals today. He worked for three years with the global library cooperative OCLC as Technology Evangelist, promoting and demystifying the benefits of Linked Data inside and outside of the organisation, to the broader library, Web and Semantic Web communities. A speaker with a global reputation for insightful, thought provoking yet understandable, and entertaining keynote sessions at major Data, Web, Semantic Web, and library focused conferences and events. Through these scene setting presentations, workshops, and facilitated panels, Richard influences the thinking and debate around these general technologies and their applicability to individual sectors and organizations.

As Chair of the W3C Schema Bib Extend working group, and an active member of the Schema.org W3C Community Group, he is at the heart of the development of the Schema.org data vocabulary that is driving the massive growth in structured data sharing on the web that underpins significant advances by Search Engines (Google, Bing, Yahoo!, Yandex) such as Semantic Search, and Knowledge Graphs.

ABSTRACT

The theme of this session is the exploration of the following aspects of Linked Data and its place in the Digital and other Library domains:

Where – did Linked Data come from

The term Linked Data was first coined by Sir Tim Berners-Lee in 2006. Although there were some enthusiasts from the world of libraries engaged from the beginning, it has only become of significant interest and a topic for discussion in our domain over the last few years. Understanding that history and its roots in early Web and Semantic Web developments, should help us understand why Linked Data is how it is and the potential benefits that flow from that.

What – is Linked Data

Not intended as a course in Linked Data, it is worth however briefly exploring its basic underlying RDF format and; how its two basic patterns of establishing unique identifiers for things then relating them to other things and their descriptive properties.

This simple standardised format, approach, and principles has led to the emergence of the Linked Open Data cloud of hundreds of fundamentally interoperable datasets across the Web. Several National Libraries actively participated by openly publishing their bibliographic data through Linked Data services.

On top of the basic foundation RDF format of Linked Data, there is a descriptive layer of vocabularies and ontologies. A combination of several well-known vocabularies, have been built upon to provide individual data models such as those used by the National Libraries and Europeana

In recent years the Library of Congress have led an initiative to produce a common vocabulary for the description of bibliographic associated resources that could be common across all libraries. BIBFRAME is still in development and as yet is to be implemented on a large scale.

Why – should we be implementing Linked Data

The adoption of Linked Data as the default way of holding data about the resources of a library has three potential general benefits – efficiency, enrichment, and enhanced discoverability.

Efficiency. In a traditional record based system, each article record for example, will contain a description of the work, the author, the publisher, the subject, etc. Whereas in a Linked Data based system there would only be one description of each person (author), each organisation (publisher), etc.

Enrichment. With focus on single descriptions of persons, organisations, and other entity types, those cataloguing can invest more in producing quality descriptions, and linking to external descriptions, knowing that the benefit will be seen wherever an entity is referenced.

Enhanced discoverability. Building discovery interfaces on Linked Data technology enables radical improvements to the discovery user experience, allowing search by any entity type, such as work, person, organisation, and for example showing lists of articles by alongside articles by an author.

Equally important is the ability to share Linked Data on the web and hopefully enhance the ability for others to find your resources.

What – should we be doing

Despite the several years that Linked Data has been around and the initiatives described, we are still at the early stages of general implementation. To move things forward and gain the benefits we need to realistic about what is holding us back.

System suppliers have been slow to introduce linked data at the core of their systems. For them it is a significant investment into a market with, as yet, little consensus as to what is needed.

BIBFRAME holds out the promise of becoming a unifying vocabulary with sufficient detail capabilities to satisfy the needs of libraries. Unfortunately, progress is slow and very Library of Congress focused – there are other initiatives to push it forward but they also have the effect of fragmenting the initiative.

One potential win would be for all to share linked data descriptions of our resources with the search engines via our web interfaces. This should increase discover ability on the web significantly.

Unfortunately, the search engines are not looking for vocabularies such as BIBFRAME. They are demanding Schema.org – the generic Linked Data vocabulary already adopted by 10s of millions of websites.

There have been efforts in sharing library data using Schema.org – OCLC added it to WorldCat in 2012 and others are in the process. Although Schema.org is up to the job of describing resources for web discovery, it is not the right tool for supporting library cataloguing practice. This disconnect between internal cataloguing and external visibility is another area which has slowed adoption in our community.

This review of the Linked Data history and landscape, will help form a view of the pragmatic next steps in applying Linked Data for real benefit.

ENNO MEIJERS

Information Manager,
Koninklijke Bibliotheek,
National Library of The Netherlands

25TH OCT, 6:00-6:40PM

Enno Meijers works as an information manager at the Koninklijke Bibliotheek, National Library of the Netherlands. His main focus is metadata management and discovery services. The National Library is one of the partners in the Digital Heritage Network (NDE). The NDE is a national program aimed at increasing the social value of the collections maintained by the Libraries, Archives and Museums in the Netherlands by improving their sustainability, usability and visibility. For the past year he has been working on the development of a new cross domain discovery infrastructure for the heritage collections. Linked Data and distributed web technologies will be at the core of this new infrastructure. Earlier he has been responsible for the development of the National Library Catalogue platform build as one of the services for the national digital infrastructure for the Public Libraries . He is one of the founders of the Dutch DBpedia chapter and currently a member of the DBpedia Association Board. He studied Electrical Engineering and Business Informatics and has been working for libraries for the past eighteen years.

A DISTRIBUTED NETWORK OF HERITAGE INFORMATION

ABSTRACT (By Enno Meijers & Sjors de Valk)

The Dutch Cultural Heritage sector is working together in a joint effort to improve the usability of its cultural heritage collection data. Our challenge is to develop a digital heritage infrastructure that overcomes the necessity of aggregating and post-process data. Instead we aim to realize a true distributed network of digital heritage information. This paper describes our approach for developing a new, cross-domain discovery infrastructure for the Dutch heritage collections. With this new infrastructure we expect to improve the usability of the collection data maintained by the heritage institutions. Implementing Linked Data principles in the collection registration systems is one of central building blocks of this approach. We urge the maintainers of the collections to align their data with formal Linked Data resources, like thesauri (people, places, periods, concepts) and to publish data as Linked Open Data.

The Dutch Digital Heritage Network (NDE) program is a national program aimed at increasing the social value of the collections maintained by the libraries, archives and museums in the Netherlands. The partners in the NDE network are the Ministry of Culture, the National Library, the National Archives, the Institute for Sound and Vision, the Cultural Heritage Agency and a number of Research Institutes for Dutch Culture and History. These partners are formalizing their co-operation through the installment of a new organization that will be responsible for realizing a joint strategy program for the Dutch cultural heritage network. The goal is a distributed network build by the institutes and their stakeholders (including commercial parties), each contributing from their own perspective. The program consists of three layers with a functional division between the management of data collections ('sustainability'), facilities for connecting that data ('usability'), and applications for presentation and use of the data ('visibility').

Our work at the usability layer is focused on the development of a lightweight cross-domain infrastructure that is built on a distributed architecture. The core functionality consists of a 'network of terms' that references all common definitions for places, people, concepts, time periods. These terms are made accessible through an SKOS API which collection registration systems can implement in order to search for relevant terms when annotating their cultural heritage objects. As a result references to formalized terms (URIs) will be added to the object descriptions. The NDE program works on getting all relevant terminology sources, maintained by the institutions, available as Linked Data and provide facilities for term alignment and even support for building new thesauri. Several tools for this work (CultuurLink [1], PoolParty [2], OpenSKOS [3]) are being provided by the NDE network.

Having cultural heritage institutions publish their data as Linked Open Data with references to established definitions for places, persons, time periods and concepts is one part of the challenge. The other part is to provide means for browsing in a cross-domain, user centric fashion. Based on possible relevant URIs identified in the user query we want to be able to browse the available linked data in the cultural heritage network. In general, the concept of 'browsable linked data' is still a challenging concept. Although Tim Berners-Lee describes the concept of 'browsable graphs' and even states that statements which relate things in two documents must be repeated in each, this is not a common practice in the Linked Data world [4].

When browsable Linked Data is offered then it is limited to the 'follow your nose' principle which is only based on using forward links. In order to really navigate in a bidirectional way through the Linked Open Data cloud, support for navigating using back links is needed as well. Most Linked Data projects make bidirectional navigation work by aggregating Linked Data dumps and loading them in a triple store where both entries to the resource can be queried. This approach of physical data integration requires data replication and large central infrastructures which is undesirable for our quest for a distributed network of digital heritage information. An alternative approach is virtual data integration using federated queries. The current Linked Data solutions (triple store with SPARQL endpoints) are hard to implement for small organizations and suffer from major performance issues when used in a federated setting. A more feasible approach is implementing the Linked Data Fragments technology as described in ' Towards sustainable publishing and querying of distributed Linked Data archives' [5]. It is our goal to implement this solution in many institutions.

But even using federated querying based on Linked Data Fragments leaves us at the question which endpoints have relevant data for a specific user question. The Dutch Digital Heritage Network consists of about 1500 institutions that hold collections. Random querying all the endpoints in this network using Linked Data Fragments would be impractical and unrealistic. A preselection of most relevant data sources relevant for the query is necessary.

These considerations have led to our decision to build a registry (preferably distributed) that records the backlinks for all the terms used in the Digital Heritage Network. The registry contains the formal Linked Data definitions of all the institutions and a high level description of the datasets. In addition to this, we will also build a 'network of relations' that describes the relations between the objects in the collection and the formalized terms used in the object description. This information provides the back links and makes it possible to navigate from a term URI to the objects that have a relation with this term.

With this new approach we hope to move away from a traditional repository centric approach to a more web centric approach where optimizing the usability of resources in their source environment is the starting point. We have published a high-level design for our distributed network of heritage information [6] which is currently being reviewed by the stakeholders. We are working on implementing parts of this this functionality in several projects and we hope to demonstrate the first results at the end of 2017.

---

[1]: http://cultuurlink.beeldengeluid.nl
[2]: https://www.poolparty.biz/
[3]: https://github.com/OpenSKOS
[4]: https://www.w3.org/DesignIssues/LinkedData.html
[5]: http://linkeddatafragments.org/publications/jod2017.pdf
[6]: https://github.com/netwerk-digitaal-erfgoed/high-level-design

Theme:

SEARCH & INFORMATION RETRIEVAL

Talk 1: Handling Noise in Scholarly Reserach: Challenges & a Case Study

Talk 2: Academic Search: Challenges and Possibilities

Talk 3: Multilingual and Crosslingual Retrieval in Digital Libraries

PANEL 2: ACADEMIC SEARCH ENGINES, DISCOVERY TOOLS AND PUBLISHERS - COLLABORATION OR COMPETITION

ANURAG ACHARYA
Distinguished Engineer,
Google Inc. USA

26TH OCT, 9:00-9:40AM

HANDLING NOISE IN SCHOLARLY SEARCH: CHALLENGES & A CASE STUDY

Anurag Acharya is a Distinguished Engineer at Google. He is one of the founders of Google Scholar. Previously, he led the indexing group at Google. He has a Bachelors in Computer Science from the Indian Institute of Technology, Kharagpur and a PhD in Computer Science from Carnegie Mellon. Prior to joining Google, he was a post-doctoral researcher at the University of Maryland, College Park and an assistant professor at the University of California, Santa Barbara.

ABSTRACT

Handling noise is one of the major obstacles for building a comprehensive scholarly search. I will describe the diverse ways noise is generated in scholarly communication and the processes that amplify noise. I will present key approaches for handling noise using Scholar Profiles as a case study.

> “Handling noise is one of the major obstacles for building a comprehensive scholarly search.”

MIHAELA VORVOREANU

Senior Program Manager,
Microsoft Research

26TH OCT, 9:40-10:20AM

Dr. Mihaela Vorvoreanu (aka: Mickey; Dr. V) is a Senior Program Manager with the Internet Services Research Center at Microsoft Research, and an Associate Professor in the Department of Computer Graphics Technology at Purdue University, where she served as graduate chair and founded an undergraduate program in UX Design. Dr. Vorvoreanu engages in use-inspired basic research in the area of democratizing knowledge. She was lead designer/researcher for an online platform for mining and visualizing National Science Foundation investments, DIA2.org. At Microsoft, she leads design, outreach, and research for Microsoft Academic - a project that aims to democratize artificial intelligence and scholarly knowledge. One of the outputs of this research project is the website http://academic.microsoft.com which empowers users to discover academic knowledge using semantic search powered by artificial intelligence.

ACADEMIC SEARCH: CHALLENGES AND POSSIBILITIES

ABSTRACT

Currently, academic search, and the search field as a whole, adequately supports users who know what they are looking for. In situations when users have a specific item or concept in mind, academic search is quite efficient at helping them locate it. The success of a search depends on users' mastery of the appropriate keywords. Often, librarians and professors are asked by searchers to help them identify the right keywords. But, what do we do when users do not have access to a librarian, a professor, or even a library? How can academic search support users who do not know exactly what they are looking for, but they have a broad idea of a topic they think they are interested in? These questions point to an unmet need in academic search. To meet this need, we need to shift the paradigm of search from supporting the location of a known item to knowledge discovery and learning. This search paradigm is known as exploratory search. This talk explains how Microsoft Academic, a service that empowers users to discover knowledge through semantic search, enables exploratory search. A few examples are provided, illustrating use cases of exploratory search supported by Microsoft Academic, that would simply not be possible in the paradigm of known-item search. The knowledge graph that powers such use cases is then explained, followed by its availability and possible uses by entities other than individual users. On a broader level, this talk identifies challenges and opportunities for academic search as well as possibilities for future growth.

> “How can academic search support users who do not know exactly what they are looking for, but they have a broad idea of a topic they think they are interested in?”

SUDESHNA SARKAR

Professor & Head,
Computer Science & Engineering,
IIT Kharagpur

26TH OCT, 10:40-11:30AM

Sudeshna Sarkar is a Professor and currently the Head of the Department of Computer Science and Engineering at IIT Kharagpur. She completed her B.Tech. in 1989 from IIT Kharagpur, MS from University of California Berkeley in 1991, and PhD from IIT Kharagpur in 1995. She served in the faculty of IIT Guwahati and at IIT Kanpur before joining IIT Kharagpur in 1998. Her research interests are in Artificial Intelligence, Machine Learning and Natural Language Processing. She has been working on Text Mining and Information Retrieval systems, and is very involved in developing natural language processing resources and tools for Bengali. She is also currently working on applications of Machine Learning in several domains such as climate science and retail.

MULTILINGUAL AND CROSSLINGUAL RETRIEVAL IN DIGITAL LIBRARIES

ABSTRACT

Digital libraries enable retrieval of information and knowledge by making books and materials available to a large audience by removing physical and spatial barriers. But universal dissemination is only possible if the materials in a digital library are truly accessible to all by removing barriers of language and culture. We live in a multilingual world having about 7000 living languages. Information too is published regularly in many of these languages. Language is a fundamental attribute of cultural identity and empowerment and digital libraries must endeavor to provide access to all in their native language to materials published in different languages. In this talk we will discuss the technological issues related to the development of multilingual digital libraries and point to technological advancements that can power them.

Crosslingual information retrieval deals with retrieving information written in a language different from the language of the user's query. Multilingual retrieval systems process information in multiple languages. They may handle queries in different languages and provide results in multiple languages. A broader term is access and we may have crosslingual and multilingual information access systems that go beyond retrieving materials to also making sure that the materials are understandable to the readers given their linguistic familiarity. Indexing and retrieval technologies must enable users to search for and discover digital materials in different languages and media, and also make way for understanding and reuse of the materials by the users.

The technologies that enable multilingual access include language identification, character encoding standards, script rendering, indexing, ranking and retrieval, and translation. Multilingual systems must use metadata that support multilingual access. The documents must be accessible to users in different languages. A digital library is supported by an information retrieval system that accepts user queries and display relevant results to the user. This system has two main components – indexing and matching.

During the indexing phase, a suitable representation of the documents and queries is used. The indexing pipeline may use different language specific components such as tokenization, normalization, stemming and multiword identification. In addition there are issues related to multiple word senses (polysemy) as well as synonymy. These need to be handled using word sense disambiguation methods and by techniques such as latent semantic indexing. A multilingual retrieval system needs to be enabled for doing these steps for multiple languages or work with language independent techniques for some of the above.

The two standard methods to support crosslingual information retrieval are query translation or document translation of which the former is more popular. However the task is challenging due to query ambiguity, translation ambiguity and insufficient lexical coverage. Dictionary coverage is typically poor for less resourced languages as well as technical terms or domain terms which are most important for querying. Apart from dictionaries multi-lingual retrieval systems can make use of parallel sentences, document-aligned as well as comparable corpora to improve the effectiveness of retrieval. In natural language processing tasks, it has now become popular to represent words as vectors in some latent space. There have been various work aimed at representing words from multiple languages in the same vector space. The use of word vectors provides a promising direction in multilingual information retrieval. In fact the use of multiple languages has been found to improve the effectiveness of some retrieval systems and to benefit those languages for which there are limited resources. Multilingual retrieval systems may use multiple indexes for different languages or a unified index containing documents from all languages.

But true access to the multilingual content requires high quality translation of results and presentation to the user. Machine translation is one of the most challenging problems for researchers. Rapid strides have been made in the last few years by making use of statistical machine translation and deep learning based techniques. Statistical machine translation has proved to be very powerful but the quality of the translation depends on the amount of parallel data available between two languages. However the research in this area has been progressing rapidly and the use of a single model for translating between multiple languages has led to dramatic improvements in translation quality involving less resourced languages. Apart from these issues, the design of an effective search interface is very important for multilingual access systems so that they are able to assist the user. Services that may be important are translation services, query suggestion, capturing feedback on query from the user and document translation assistance.

While challenges still remain, the advancements in natural language processing should be harnessed for building multilingual digital libraries.

> “The technologies that enable multilingual access include language identification, character encoding standards, script rendering, indexing, ranking and retrieval, and translation.”

Panel 2:

Academic search engines, discovery tools and publishers: collaboration or competition

26th October
11:30AM - 1:00PM

The Panelists

MIHAELA VORVOREANU
Microsoft Research

PAUL D CLOUGH
University of Sheffield

BALAJI DEVARAJAN JAYARAMAN
Taylor & Francis Group

PAUL NIEUWENHUYSEN
Vrije Universiteit Brussel

JOHN AKEROYD
CIBER Research

MODERATOR:
DR. JAGDISH ARORA
INFLIBNET

Theme:

DESIGNING FOR THE USERS

Talk 1: Making Users Stakeholders – Strategies to Engage and Collaborate

Talk 2: User-informed Design for Search and Discovery in Digital Culture Heritage

Talk 3: Digital Library and Accessibility

Talk 4: OCR for Indian Languages and Beyond

Talk 5: Crowdsourcing Based Curation and User Engagement in Digital Library Design

Talk 6: Online Learning and Open Educational Resources: Challenge of the Future (with a glimpse at the past)

JILL COUSINS
Executive Director,
Europeana Foundation

26TH OCT, 2:00-2:40PM

Jill Cousins is the Executive Director of Europeana Foundation, building up Europeana from a project idea to an operational service. She has many years' experience in web publishing. Her past experience includes the commercial publishing world as European Business Development Director of VNU New Media and scholarly publishing with Blackwell Publishing running their online journals service. Prior to publishing she had a variety of marketing and research careers in the information field. All are very divergent from her academic career in historical cartography and a Ph.D in Sixteenth Century Arabic and Turkish Sea Charts

MAKING USERS STAKEHOLDERS: STRATEGIES TO ENGAGE AND COLLABORATE

ABSTRACT

Reading, listening, watching and viewing are superficially passive activities. The writer writes, I read, a song is played, I listen, a picture painted, I look. But this belies human nature which is to question, interpret, analyse. The difference is that nowadays we have the technologies to facilitate these reactions immediately and possibly more importantly in the context of this talk, users expect to be able to contribute. This is enormously powerful, not only does it help us deliver what people are actually looking for, but gives us a huge creative, problem-solving capacity where our audience become our partners and collaborators – true stakeholders.

This talk will cover Europeana's overall strategy to engage and encourage participation with users. This has 3 main strands and is linked to two elements of our overall strategy[1] scaling with partners and engage people.

Scaling with partners: We know that we cannot be all things to all men. We don't have the capacity or knowledge, but we do know people who do. Using this philosophy we have partners in Education, in Research and in the Creative Industries; people who can use our material and make it work for their audiences. This has led to a multitude of collaborations and scalability. **Engage people:** getting people to the thing they are looking for in massive amount of data in multiple languages and media types that is Europeana cannot be solved by clever search and discovery technologies alone. We have therefore created audience or subject specific sites (europeana collections)[2] where we can for instance engage researchers in data, showcasing their thoughts and interactions such as the results of our research grants programme[3].

We also place the content in peoples workflow, working with third party platforms (such as Wikipedia, Pinterest, Facebook) to engage and collaborate and develop real user participation and collaboration activities to improve the data and harness the power of the user as a collaborator and contributor. This includes our Europeana1914-1918.eu Collection days, Competitions, Transcribathons, Picture This, Wiki loves Public Art etc.

We can and should be doing even more to make sure that our audience are real stakeholders in the digital library, improving, solving and participating.

1 https://strategy2020.europeana.eu
2 https://europeana.eu
3 https://pro.europeana.eu/post/europeana-research-grants-2016-outcomes-thylstrup

PROF. PAUL D CLOUGH

Professor of Information Retrieval, University of Sheffield, UK

26TH OCT, 2:40-3:20PM

USER-INFORMED DESIGN FOR SEARCH AND DISCOVERY IN DIGITAL CULTURAL HERITAGE

Paul Clough is Professor of Information Retrieval at the Information School, University of Sheffield. He received a B.Eng. (hons) degree in Computer Science from the University of York in 1998 and a Ph.D. from the University of Sheffield in 2002. Prior to joining Sheffield he worked for British Telecommunications Plc. His research interests mainly revolve around developing technologies to assist people with accessing and managing information. In particular, Paul has published work in the areas of multilingual information retrieval, information access to digital cultural heritage, evaluation of search systems, geographical information retrieval, text-based image retrieval, plagiarism detection, text re-use, and search analytics. Paul was Scientific Director for the FP7 EU-funded Personalised Access to Cultural Heritage Spaces

ABSTRACT

This talk will consider the importance of designing systems and services to support users in their search and discovery activities. By having a better understanding of the types of users engaging with digital cultural heritage and their goals and tasks enables us to take a more user-centred focus to the design and evaluation activities. This is becoming increasingly important for the providers of cultural heritage as they seek to make digital content discoverable and accessible, go beyond simple document lookup and support exploratory searching activities, and seek to provide rich user experiences to a broad range of user groups, from experts to the general public. I will draw upon past and current research projects I have been involved with organisations, such as the UK National Archives, National Museums Liverpool and Europeana, to help illustrate the value of understanding users and the context in which they are using digital cultural heritage. In particular the talk will also discuss two studies conducted recently and presented at the TPDL 2017 conference in Thessaloniki:

(i) a study of the search tasks carried out by users of Europeana and their uses of information; and

(ii) a study of "general users" of the website at National Museums Liverpool.

> “The technologies that enable multilingual access include language identification, character encoding standards, script rendering, indexing, ranking and retrieval, and translation.”

PROF. ANUPAM BASU

Chairman, Centre for Educational Technology, Professor, Computer Science & Engineering, IIT Kharagpur

26TH OCT, 3:20-4:00PM

Prof. Anupam Basu, Professor, Dept. of Computer Science & Engineering, IIT Kharagpur, is an active researcher in the areas of Cognitive and Intelligent Systems, Embedded Systems and Language Processing. Presently he is the Executive Director of the Indian Institute of Information Technology, (IIIT) at Kalyani and also the Chairman and Head of the Center for Educational Technology, IIT Kharagpur. He leads a number of projects for Technology Enabled Education, specially for School Education.

He is an Alexander von Humboldt Fellow and a Fellow of the Indian National Academy of Engineering. The awards won by him include the State Award for the Best Contribution to the Cause of Empowerment of the Disabled (2014), Universal Design Award 2011, for contributions in design for the disabled, by National Council for Promotion of Employment of Disabled Persons, India, the National Award for the Best Technology Innovation for the Physically Disabled (2007) and the Da-Vinci Award 2004 from the Engineering Society of Detroit

DIGITAL LIBRARY AND ACCESSIBILITY

ABSTRACT

While the advances in the digital world are revolutionizing the ways of living and learning, it is often overlooked that it is creating digital divides to varying degrees for different segments of the population. The systems, tools and the novel applications are addressing themselves to several of the developmental needs of the modern society, yet the novelties are most often deprived of the accessibility features that enables "all" to use them with equal ease. The variety of disabilities, physical, cognitive as well as geriatric, calls for adopting a "universal design" approach to be adopted at the very onset of the design phase. The systems to deliver digital contents over the internet and through digital libraries face the same requirement and challenge. While international standards are being proposed and being gradually adopted, incorporation of novel features at a later stage also push in non-compliance to such standards.

In this talk, we will consider some of the demands and requirements posed by the disabled community to the delivery of the digital contents and will dwell on some of the measures that should be undertaken at the development of content phase as well as by the delivery interfaces of digital libraries. In particular, we will touch upon video as well as textual web contents.

"The systems, tools and the novel applications are addressing themselves to several of the developmental needs of the modern society, yet the novelties are most often deprived of the accessibility features that enables "all" to use them with equal ease."

PROF. C V JAWAHAR

Professor & Amazon Chair Professor, International Institute of Information Technology, Hyderabad

26TH OCT, 4:20-5:00PM

OCR FOR INDIAN LANGUAGES AND BEYOND

C. V. Jawahar is a professor at IIIT Hyderabad, India. He received his PhD from IIT Kharagpur and has been with IIIT Hyderabad since Dec. 2000. At IIIT Hyderabad, Jawahar leads a group focusing on computer vision, machine learning and multimedia systems. In the recent years, he has been looking into a set of problems that overlap with vision, language and text. He is also interested in large scale multimedia systems with special focus on retrieval. He has more than 50 publications in top tier conferences in computer vision, robotics and document image processing. He has served as a chair for previous editions of ACCV, WACV, IJCAI and ICVGIP. Presently, he is an area editor of CVIU and an associate editor of IEEE PAMI. He is also a program co-chair for ICDAR 2017 and ACCV 2018.

ABSTRACT

OCRs allow the content level access to the scanned or digitized document collections. However, these technologies are not yet fully available for Indian languages for deployment. In this talk, we discuss the associated challenges, state of the art and the recent developments.

In addition to the conventional OCRs, we also discuss a set of associated developments in recognition and retrieval of document images. Recent advances in machine learning and specially deep learning has generated a new promising direction. This talk will conclude by discussing the missing links in establishing reliable OCR systems for Indic scripts for public use.

“OCRs allow the content level access to the scanned or digitized document collections.”

CROWDSOURCING BASED CURATION AND USER ENGAGEMENT IN DIGITAL LIBRARY DESIGN

ROSE HOLLEY

Special Collections & Digital Curator,
University of New South Wales, Canberra

26TH OCT, 5:00-5:40PM

Rose Holley has been working as professional librarian in the UK, NZ and Australia for more than 30 years and is passionate about client focussed services. Since 2014 Rose has held the position of Special Collections and Digital Curator at UNSW Canberra. Prior to this she held senior positions at the National Archives of Australia (Director of Software Implementation 2012-2014), the National Library of Australia (Director of Trove and the Australian Newspapers Digitisation Program 2007-2012), and the University of Auckland (Manager of Digital Services and Projects 2000-2007). She has expertise in collection management, digitisation projects, digital preservation and innovation. She is an author for several library journals and actively contributes to library and archival professional associations, committees and groups. Her area of research is public crowdsourcing for archive and library collections.

ABSTRACT

It is now 10 years since the first mass public crowdsourcing project was started in a library. This was the innovative text correction of historic Australian Newspapers for the National Library of Australia. It was a phenomenal and unexpected success, more initially by chance, novelty and experimentation, than considered design. Back in 2007 very little was known about crowdsourcing behaviours, and Wikipedia was still relatively new. The decision to implement public text correction and encourage our users to improve and add value to library content was initially driven by a desire to have a high quality search, rather than any wish for social engagement or online community building. The idea of enabling anyone to alter or 'curate' library owned content was terrifying and controversial for some librarians, whilst the users mostly saw it as a logical and indeed very overdue step in their expectations of how they should be able to engage with content in a digital world.

Over the last ten years I have undertaken research into the behaviours, motivations and needs of both the 40,000 public newspaper volunteers who participate in crowdsourcing and the libraries who offer crowdsourcing features on their content, to better understand what is happening and why. I have compared the activity of Australian Newspapers to other significant but similar non-library services with large communities of online volunteers such as Wikipedia and Zooniverse. I've also looked at what works and what doesn't on user interface design and how to be totally user centric when designing these services, working in partnership with your users, or even encouraging them to lead you, as we did in the first three years of Australian Newspapers.

In this session I will share the knowledge I have gained on crowdsourcing, looking at a variety of library examples, together with my recommendations. We will take an in depth look at the Australian Newspaper Text Correction aspect of Trove over the last ten years, including key decision points in the crowdsourcing journey; management and encouragement of the online volunteers and their achievements. I also look to the future and question if we were doing it all again, starting now, would we do it the same or differently for correction of newspapers? With artificial intelligence and machine learning rapidly advancing, and research ongoing in post OCR text correction programs through artificial intelligence training (now that significant corpus's of both machine corrected and manual corrected newspaper text are available), is it likely that crowdsourcing will continue, or will it be replaced by implementation of smarter intelligent software programs? Where do humans add value vs machines? Also when we say it is 'successful' what do we really mean?

As our digitisation outputs increase through digitisation of millions more pages, our volunteers are faced with massive amounts of uncorrected text, and their phenomenal accomplishments are but a drop in the ocean.

If the social engagement aspect of crowdsourcing is removed from our digital libraries, then in what other ways can we encourage user engagement? I will discuss some examples of virtual and physical community based learning focus groups that arose from the use of Australian digitised newspapers, such as climatologists, knitting enthusiasts, medical research, language research, local history groups, natural disaster support groups, and transport enthusiasts.

In the end what is the cost vs the benefit of crowdsourcing and its future? How can we leverage crowdsourcing to our best advantage, applying what we have learnt in the last ten years?

> "The idea of enabling anyone to alter or 'curate' library owned content was terrifying and controversial for some librarians, whilst the users mostly saw it as a logical and indeed very overdue step in their expectations of how they should be able to engage with content in a digital world."

ONLINE LEARNING AND OPEN EDUCATIONAL RESOURCES: CHALLENGE OF THE FUTURE (WITH A GLIMPSE AT THE PAST)

PROF. RORY McGREAL

Professor, Centre for Distance Education, Athabasca University, Alberta, Canada

26TH OCT, 5:40-6:20PM

Rory McGreal is a Professor in the Centre for Distance Education at Athabasca University AU), Alberta Canada. He is also the UNESCO/Commonwealth of Learning/International Council for Open and Distance Education Chair in Open Educational Resources and Director of the Technology Enhanced Knowledge Research Institute (TEKRI). In addition, he is the co-Editor of Canada's first open access journal, The International Review of Research in Open and Distributed Learning (IRRODL). He was the Associate Vice President, Research at AU for ten years. Previously, he was the executive director of TeleEducation New Brunswick , a province-wide bilingual (French/English) distributed distance learning network. Before that, he was responsible for the expansion of Contact North (a distance education network in Northern Ontario) into the high schools of the region. His Ph.D. degree (1999) in Computer Technology in Education at Nova Southeastern University 's School for Computer and Information Science was taken at a distance using the Internet. Rory was the founder of the world's first e-learning website for TeleEducation NB and one of the world's first metadata learning object repositories, the TeleCampus. In the past, he has worked in Canada as a teacher and teacher representative, and abroad in the Seychelles , the Middle East and Europe in various capacities as a teacher, union president, ESL technological training co-ordinator, instructional designer, language and computer laboratory co-ordinator, and educational advisor.

Awards:

European Distance Education Network (EDEN) Senior Fellow October, 2016

Canadian Network for Leadership in Education Award, from the Canadian Network for Innovation in Education, Waterloo, ON, May, 2016

Lifetime Achievement Award, Open Education Consortium, Krakow, Poland, April, 2016

Open Courseware Consortium Award for the OER Knowledge Cloud 2014

ABSTRACT

Higher education institutions worldwide continue to face significant challenges related to providing increased access to high quality education, while containing or reducing costs. New developments in higher education all speak to the efforts on the part of the traditional higher education community, as well as more flexible providers such as open universities, the traditional higher education community, as well as more flexible providers such as open universities, to address these challenges. Such developments have the potential to increase access and flexibility in higher education. Basic education for all continues to be a goal that challenges – and will continue to challenge – many countries.

Open Educational Resources (OER) constitute an important resource with the potential to facilitate the expansion of quality education and learning opportunities worldwide. OER refers to full courses, course materials, modules, textbooks, streaming videos, tests, software, and any other tools, materials or techniques used to support access to knowledge. The free and open sharing of educational resources can serve to promote the building of knowledge societies and the reduction of the knowledge divide that separates nations, as well as the divide within societies themselves.

The relevance of OER is augmented by the exponential growth in online accessibility afforded by the wide range of new applications and devices. Moreover, this is happening at an increasingly rapid pace. Security, privacy and consistency coupled with user control. The growing trend toward online learning using the power of networks has opened the door for learners and teachers to access the world's knowledge from almost anywhere, at anytime. The internet houses the world's treasure of knowledge. It is the world's intellectual commons that renders knowledge accessible to all. The world's knowledge is a public good that should be made available to everyone.

Theme:
ACCESS & RIGHTS

Talk 1: Rethinking the Future of Openness

Talk 2: Delocalising Knowledge..... IPR Issues Hold the Keys

Talk 3: Do the Right Thing: Interoperable Rights Statements

Talk 4: Preservation of Documentary Heritage: UNESCO Experience

PANEL 3: NETWORKING AND COLLABORATION AMONG DIGITAL LIBRARIES OF THE WORLD

PROF. LAWRENCE LIANG

Dean,
School of Law, Governance and Citizenship,
Ambedkar University Delhi

27TH OCT, 9:30-10:10AM

Lawrence Liang is the Dean, School of Law, Governance and Citizenship, Ambedkar University Delhi. Prior to AUD Lawrence was a cow founder of Alternative Law Forum, a pubic interest lawyering group based in Bangalore and has taught in several institutions in India, the US and Turkey. Lawrence works on intellectual property and public interest and was involved in the DU photocopy case and has served on the Board of the I Commons. He was also formerly the legal lead for Creative Commons.

RETHINKING THE FUTURE OF OPENNESS

ABSTRACT

The Creative Commons has been around for at least fifteen years now and has made impressive inroads both into the normative discourse of copyright as well as providing an alternative licensing mechanism for a number of knowledge/ culture initiatives. And yet there seems to be relatively little purchase of open content licensing outside a small circle of activists. In this presentation I will take stock both as an insider to the CC system as well as a sympathetic critic of it to ask examine the challenges we face while translating the politics and form of Open content movements into countries like India. I will also examine the limitations of such initiatives developing an argument of the distinction between genteel and carnal openness.

> “The Creative Commons has been around for at least fifteen years now… And yet there seems to be relatively little purchase of open content licensing outside a small circle of activists.”

DELOCALISING KNOWLEDGE..... IPR ISSUES HOLD THE KEYS

PRABUDDHA GANGULI

CEO, VISION-IPR, Mumbai
& Visiting Professor,
Rajiv Gandhi School of Intellectual Property Law, IIT Kharagpur

27TH OCT, 10:10-10:50AM

He obtained his B.Sc. degree in Chemistry from Bombay University, M.Sc in Chemistry from Indian Institute of Technology, Kanpur and PhD in Chemistry from the Bombay University based on his work at the Chemical Physics Department, Tata Institute of Fundamental Research (TIFR). This was followed by postdoctoral research in Germany at the Johannes Gutenberg Universitat, Mainz as an Alexander von Humboldt Foundation Fellow and at Windsor University, Canada.

He held the Ministry of Human Resources Development (MHRD) IPR Chair Professorship at the Tezpur University, Assam, India from May 2012 to July 2017. He has been a member of the Board of Management of the Autonomous Physics Department of the University of Mumbai. He is Honorary Scientific Consultant for Innovation and IPR matters to the Office of the Principal Scientific Adviser, Government of India, a member of the Advisory Board of the IIPI, Washington and member of the International Editorial Board of "World Patent Information. Since July 2017, he is also Visiting Professor at the Rajiv Gandhi School of Intellectual Property Law in the Indian Institute of Technology, Kharagpur.

Since 2001 he is CEO of his consulting firm as CEO of VISION-IPR offering services in Intellectual Property Rights and knowledge management. He is a practicing Patent Agent at the Indian Patent Office and is a strategic IPR consultant to several companies in India and abroad helping them to integrate IPR in their business processes for facile translation of innovations into markets. As a WIPO Consultant, he has been an adviser to several governments on national innovation and IPR policy matters including participating as trainer in their IPR capacity building programmes.

In February 2011 he was awarded "The Chemtech Pharma-Bio World Award for outstanding contribution in the field of Intellectual Property". In 2014, the legal fraternity in India under the banner of "LEGAL ERA" honoured him with "The Lifetime Achievement Award" for his continued contribution in the field of IPR. In October 2014, The Tata Institute of Fundamental Research (TIFR) Alumni Association conferred on him the "Excellence Award" for his contributions in the field of Technology Management and IPR.

ABSTRACT

Digital Technologies have continually metamorphosed our society as we experience today and is set on an autonomous transformational trajectory with infusion of novel man-machine interactive processes leading to delocalisation of knowledge as a necessary and unstoppable consequence. The "digital-collective" way of life is a societal reality that will dictate the way in which we unshackle the barriers to knowledge to propagate the promise of "Open Education" exploiting global information resources, knowledge and expertise. Creative and user friendly "Digital Libraries" [DLs] are expected to facilitate effective linking of and borderless access to the universal complex knowledge chain.

It is in this context that "Intellectual Property Rights" (IPR) becomes an inseparable partner in the "value added knowledge creation-linking-access" implementation through DLs. Utilisation of technologies that are patented, creations that are copyrighted, works that are secured with "digital technologies", distinctive marks that are trademarked, domain names that are proprietary, use of meta-tags, etc., are a few examples of IPR issues that haunt DLs with the fear of IPR infringement.

The turf involving IPR is becoming murkier with fierce debates on Intellectual Property [IP] protection, enforcement and affordability, IP and fair benefit sharing with its creators and owners, individual & collective licensing agreements and further establishing pragmatic frameworks for "Open Access".

The role of International and National Agencies is crucial in influencing IPR laws and policies in diverse jurisdictions, setting up workable IPR guidelines for DLs, training personnel in DLs to ensure compliance to set standards and design programmes to enhance awareness on how to integrate IP into functioning of DLs.

The presentation will dwell on a set of imperatives that need to be weaved into DLs practices to ensure law abiding practices in its multifaceted knowledge delocalisation process.

> "The "digital-collective" way of life is a societal reality that will dictate the way in which we unshackle the barriers to knowledge to propagate the promise of "Open Education" exploiting global information resources, knowledge and expertise."

JILL COUSINS
Executive Director,
Europeana Foundation

27TH OCT, 11:10-11:50AM

Jill Cousins is the Executive Director of Europeana Foundation, building up Europeana from a project idea to an operational service. She has many years' experience in web publishing. Her past experience includes the commercial publishing world as European Business Development Director of VNU New Media and scholarly publishing with Blackwell Publishing running their online journals service. Prior to publishing she had a variety of marketing and research careers in the information field. All very divergent from her academic career in historical cartography and a Ph.D in Sixteenth Century Arabic and Turkish Sea Charts.

ABSTRACT

In just two years the aspiration of internationally interoperable rights statements had grown from seed into a fully operational service.

DO THE RIGHT THING: INTEROPERABLE RIGHTS STATEMENTS

'RightsStatements.org establishes the vocabulary that every organization can use to talk to their audiences about copyright and related rights in a meaningful way.'

The concept of RightsStatements.org was that everyone should know how to legally engage with their cultural heritage online. The building blocks were initiated in 2014 by representatives and partners of the DPLA and Europeana. The conceptual development had two work streams managed by the Technical and the Statements Working Groups. Their work was validated in 2015 through a public consultation based on two white papers, establishing the need for such statements, the statements themselves[1] and the technical requirements for implementing them.[2]

The validated papers established that the principles of the rights statements should be: simple, flexible, descriptive, accurate and transparent. Furthermore, they should serve;

- Cultural institutions to document their best evidence of the copyright status and reuse restrictions associated with a digital object in their collection,
- Users by communicating the copyright and related restrictions associated with a digital object they wish to use.

In 2016, rightsstatements.org published twelve rights statements and usage guidelines. The statements are clustered into three categories; In Copyright, No Copyright, Unknown.

In April 2016 rightsstatements.org was launched at the DPLA fest in Washington DC. This marked a milestone for Europeana and the Digital Public library of America, founding partners of the rightsstatements.org consortium.

The membership has expanded with The Library & Archives Canada, The National Digital Library, India and Trove, Australia, who are part of the Steering Committee and governance of RightsStatements.org and we are hopeful to expand this to other continental and country libraries.

This talk will run through the need, purpose, application and use of rights statements, aiming to convince everyone to join up!

1 http://rightsstatements.org/files/160208recommendations_for_standardized_international_rights_statements_v1.1.pdf
2 http://rightsstatements.org/files/170106requirements_for_the_technical_infrastructure_for_standardized_international_rights_statements_v1.2.pdf

DR. INDRAJIT BANERJEE

Director, Knowledge Societies Division, Communication & Information Sector, UNESCO

27TH OCT, 11:50AM-1:00PM

Dr Indrajit Banerjee is the Director, Knowledge Societies Division, Communication and Information Sector at UNESCO Headquarters in Paris. He oversees UNESCO's global programmes on ICTs and access to information. He joined the organization in 2009 as Chief of ICTs in Education, Science and Culture, a section of the Communication and Information sector. Before joining UNESCO, Dr Banerjee was the Secretary-General of the Asian Media and Information Centre (AMIC) in Singapore from 2004–2009; and an Associate Professor at the Nanyang Technological University of Singapore from 2001–2009. Dr Banerjee holds a PhD with Distinction in communication studies from the Sorbonne University in Paris. He is a member of the International Communication Association (ICA), the International Association for Media and Communication Research (IAMCR) and the Association for Education in Journalism and Mass Communication (AEJMC). He has edited numerous books on ICTs and communication, and has published articles in the world's leading journals on media and communications.

PRESERVATION OF DOCUMENTARY HERITAGE: UNESCO EXPERIENCE

ABSTRACT

Dr Indrajit Banerjee's presentation will highlight some challenges associated with preserving documentary heritage within the context of UNESCO's Memory of the World (MoW) Programme, while signalling opportunities presented by digitization. It will contextualise these challenges within the framework of the UNESCO 2015 Recommendation Concerning the Preservation of and Access to Documentary Heritage, Including in Digital Form.

It will conclude by presenting three scenarios in which the National Digital Library and other documentary heritage holders could appropriate the UNESCO Recommendation in terms of three issues: (i) Enhancing public awareness of documentary heritage; (ii) Collaborating for policy advocacy; and (iii) Linking documentary heritage to the Sustainable Development Goals (SDGs).

> "The Knowledge Societies Division (KSD) is responsible for the coordination of UNESCO's overall contribution to the follow-up to the World Summit on the Information Society (WSIS)."

Panel 3:

Networking and Collaboration among Digital Libraries of the World

27th October
2:00PM - 4:00PM

The Panelists

JILL COUSINS
Europeana Foundation

PARTHA PRATIM DAS
National Digital Library of India

BRENDON McKINLEY
TROVE, Australia

DOUWE DRIJFHOUT
National Library of South Africa

UPENDRA PRASAD MAINALI
Nepal National Library

MEZBAH-UL-ISLAM
University of Dhaka

MODERATOR:
DR. RAMESH C GAUR
Jawaharlal Nehru University (JNU)

About UNESCO

United Nations
Educational, Scientific and
Cultural Organization

Founded in 1945, the United Nations Educational, Scientific and Cultural Organization (UNESCO) is a UN specialized agency responsible for coordinating international cooperation in education, science, culture and communication. It strengthens the ties between nations and societies, and mobilizes the wider public so that each child and citizen:

(a) Has access to quality education, a basic human right and an indispensable prerequisite for sustainable development;

(b) May grow and live in a cultural environment rich in diversity and dialogue, where heritage serves as a bridge between generations and peoples;

(c) Can fully benefit from scientific advances; and

(d) Can enjoy full freedom of expression, the basis of democracy, development and human dignity. UNESCO's messages are of increasing importance today, in a globalized world where interconnections and diversity must serve as opportunities to build peace in the minds of men and women.

About National Digital Library of India

National
Digital Library
of India

The National Digital Library of India (NDLI) is an integration platform for schools, colleges, universities, teachers, students, lecturers, differentially abled pupils and anybody who has a willingness to learn. We hope to facilitate e-learning for students everywhere with the help of technologically enhanced learning design and the support of Digital India. The National Digital Library of India has been developed by the Indian Institute of Technology, Kharagpur and is sponsored by the Ministry of Human Resources development (MHRD) under the NMEICT to establish a platform which will successfully bridge the divide between the vision and reality of a truly Digital India.

About our PARTNERS:

Indian Institute of Technology Delhi is a institute of national importance and one of the premier institutes of the country. One of the first 5 IITs to be established, it was established in 1961 with a vision to contribute to India and the World through excellence in scientific and technical education and research; to serve as a valuable resource for industry and society; and remain a source of pride for all Indians.

Jawaharlal Nehru University established in 1966, is one of the foremost university in India, and a world-renowned centre for teaching and research. It was Ranked number one in India by the National Assessment and Accreditation Council (NAAC) and was ranked no 2 among all universities in India by the National Institutional Ranking Framework, Government of India, in 2017.

The Energy and Resources Institute (TERI) is a research institute based in New Delhi that conducts research work in the fields of energy, environment and sustainable development. Established in 1974, it was formerly known as Tata Energy and Resource Institute and as the scope of its activities widened, it was renamed The Energy and Resources Institute in 2003.

The Indira Gandhi National Open University, known as IGNOU, is a distance learning national university located in New Delhi, India. IGNOU was founded to serve the Indian population by means of distance and open education, providing higher education opportunities particularly to the disadvantaged segments of society. It also aims to encourage, coordinate and set standards for distance and open education in India, and to strengthen the human resources of India through education. It also acts as a national resource centre, and serves to promote and maintain standards of distance education in India.

INFLIBNET Centre (Information and Library Network Centre) is an autonomous Inter-University Centre of the University Grants Commission (UGC) of India under Ministry of HRD (MHRD) located in Gandhinagar, Gujarat. INFLIBNET is involved in modernizing university libraries in India and connecting them as well as information centres in the country through a nationwide high speed data network using the state-of-art technologies for the optimum utilization of information. INFLIBNET is set out to be a major player in promoting scholarly communication among academicians and researchers in India.

UNESCO NEW DELHI CLUSTER OFFICE FOR BANGLADESH, BHUTAN, INDIA, MALDIVES, NEPAL AND SRI LANKA

1 San Martin Marg, Chanakyapuri,
New Delhi - 110021
Tel: +91 11 2611 1867/9
Fax: +91 11 2611 1861/2
Email: newdelhi@unesco.org

NATIONAL DIGITAL LIBRARY OF INDIA

Central Library • Indian Institute of Technology,
Kharagpur, West Bengal - 721302
Tel: +91-3222-282435
Email: ndl-support@iitkgp.ac.in

@NDLIndia