

IIPC Web Archiving Conference 2019

Zagreb, Croatia

6 – 7 June 2019



15 YEARS HAW

Hrvatski arhiv weba
Croatian Web Archive

Schedule

Thursday, 6 June 2019

8:00 Registration/Coffee

PLENARY (Panorama)

8:45 **OPENING REMARKS:** **Tatijana Petrić**, Director General of the National and University Library in Zagreb

9:00 **KEYNOTE:** **Harry Verwayen**, Europeana
Mirror World

Introduced and chaired by Jasenka Zajec

10:00 **LIGHTNING TALKS** chaired by *Jason Webber*: **Ji Shiyan & Zhao Danyang**: The key technologies of web information preservation and service system platform ~ **Ben O'Brien & Jeffrey Van Der Hoeven**: Technical uplift of the Web Curator Tool ~ **Helena Byrne**: From the sidelines to the archived web: what are the most annoying football phrases in the UK? ~ **Andrej Bizík, Peter Hausleitner & Jana Matúšková**: Archiving and LTP of websites and born digital documents in the Slovak Republic ~ **Sebastian Nagel**: Accessing to WARC files via SQL

10:30 Break

Track 1: Panorama

Track 2: Opera

Academic contexts

Chair: Alex Thurman

Machine learning projects

Chair: Draženko Celjak

11:00 **Tolulope Balogun & Trywell Kalusopa**: Developing a framework for web archiving of indigenous knowledge systems in selected academic institutional repositories in Africa

Mark Phillips: Leveraging machine learning to extract content-rich publications from web archives

11:20 **Dunja Majstorović & Karolina Holub**: Knowledge, habits and prospects of the use of web archives among the academic community in Croatia

Jefferson Bailey & Maria Praetzelis: From open access to perpetual access: archiving web-published scholarship

11:40 **Joe Carrano**: From the foundational web to founding a web archives: creating a formalized Web Archiving Program at the MIT Libraries

Fernando Melo: Searching images from the past with Arquivo.pt

12:00 **Bartłomiej Konopa**: Studying the past Web in Poland – current state and perspectives

Sara Elshobaky & Youssef Eldakar: Identifying Egyptian Arabic websites using machine learning during a web crawl

12:20 Q&A

Q&A

12:30 Lunch

Track 1: Panorama**Panel: National webs***Chair: Ditte Laursen***Track 2: Opera****Panel: Web archiving in Southeast Europe***Chair: Karolina Holub***Track 3: Overture****Workshop**

13:30 **Niels Brügger, Ditte Laursen, Friedel Geeraert, Kees Teszelszky, Valérie Schafer & Daniel Gomes:** Opportunities and challenges in collecting and studying national webs

Janko Klasinc, Dragana Milunović, Nemanja Kalezić, Inge Rudomino, Tamara Horvat Klemen, Žarko Mirković & Lidija Popović: Past, present and future of web archiving in Southeast Europe: experiences of Croatia, Montenegro, Serbia and Slovenia

Julien Nioche: Introduction to web crawling with StormCrawler

National web archives*Chair: Márton Nemeth***Research use***Chair: Olga Holownia*

14:30 **Ivy Lee Huey Shin & Charles Wijaya:** Sharing by the National Library Singapore on the journey towards collecting digital materials

Nick Ruest & Ian Milligan: See a little Warclight: building an open-source web archive portal with project blacklight

14:50 **Friedel Geeraert & Sébastien Soyez:** The first steps towards a Belgian web archive

Ditte Laursen & Niels Brügger: A national Web Trend Index based on national web archives

15:10 **Corey Davis, Carole Gagné & Nicholas Worby:** True North: the current state of web archiving in Canada

Jason Webber: Using secondary datasets for researchers under a legal deposit framework

15:30 **Q&A**

Q&A

15:40 **Break**

Track 1: Panorama**Track 2: Opera****Panel: Building (and rebuilding) web archiving programs***Chair: Tom Storrar***Panel: Archiving video content***Chair: Yves Maurer*

16:10 **Lori Donovan, Maria Ryan, Renate Hannemann & Garth Stewart:** Transition in the time of web archiving: building (and rebuilding) web archiving programs

Sara Aubry, Géraldine Camile & Thomas Drugeon: From videos to channels: archiving video content on the web

Workshops

17:10 **Sara Aubry & Kristinn Sigurðsson:** WARC format

Helena Byrne & Carlos Rarugal: Reflecting on how we train new starters in web archiving

18:10 **Finish**

20:00 **Conference Dinner, Restaurant Museum (Museum of Arts and Crafts)**

IIPC Web Archiving Conference 2019

Zagreb, Croatia

6 – 7 June 2019



Schedule

Friday, 7 June 2019

8:30 **Arrival/Coffee**

PLENARY (Panorama)

8:45 **Welcome to day 2**

9:00 **KEYNOTE: Dr. Anat Ben-David**, Open University of Israel

Web archives as memoryware: critical reflections on sources and methods for web history

Introduced and chaired by Mark Phillips

10:00 **DROP-IN TALKS**

10:30 **Break**

Track 1: Panorama

Track 2: Opera

Curatorial strategies

Chair: Pamela Graham

Technical infrastructure

Chair: Lauren Ko

11:00 **Sumitra Duncan & Lori Donovan**: Advancing art libraries: developing a collaborative national network for curated web archives

Miroslav Milinović & Draženko Celjak: Technical development of Croatian Web Archive in past 15 years

11:20 **Sabine Schostag**: Why, what, when and how: curatorial decisions and the importance of their documentation

Márton Németh & László Drótos: Metadata supported full-text search in a web archive

11:40 **Jessica Cebra**: Describing web archives: a standard with an identity crisis?

Rafael Gieschke & Klaus Rechert: Preserving web servers

12:00 **Lorenz Widmaier**: Divide-and-conquer – artistic strategies to curate the web

Gil Hoggarth: The infrastructure behind web archiving at scale

12:20 **Q&A**

Q&A

12:30 **Lunch**

Panel: Access policies

Chair: Nicola Bingham

Panel: Planning for sustainability

Chair: Jason Kovari

13:30 **Abbie Grotke, Alexandre Chautemps, Nicola Bingham, Maria Ryan, Alex Thurman & Daniel Gomes**: Access policies, challenges and approaches

Sylvain Bélanger, Nick Ruest, Ian Milligan & Anna Perricci (via video conference): Digital preservation strategies, the Archives Unleashed Cloud Project and Webrecorder

Collecting themes and formats

Chair: Maria Ryan

Crawling strategies and tools

Chair: Kristinn Sigurðsson

14:30 **Tobias Beinert, Markus Eckl & Florence Reiter**: Archiving and analysing elections: how can web archiving, digital humanities and political science go together?

Julien Nioche & Sebastian Nagel: StormCrawler at Common Crawl and other use cases

14:50 **Lynda Clark:** Emerging formats: discovering and collecting contemporary British interactive fiction

15:10 **Radovan Vrana & Inge Rudomino:** Croatian web portals: from obscurity to maturity

15:30 Q&A

15:40 Break

Track 1: Panorama

Training & impact

Chair: Inge Rudomino

16:10 **Daniel Gomes:** Arquivo.pt – memorial and other goodies

16:30 **Maria Praetzelis, Sharon McMeekin & Abbie Grotke:** Building the IIPC training program

16:50 **Julie Fukuyama & Simon Tanner:** Developing impact assessment indicators – making a proposal for the UK Web Archive

17:10 **Ricardo Basilio & Daniel Bicho:** Librarians as web curators in universities

17:30 Q&A

17:40 Closing remarks (Panorama)

18:00 Finish

Eléonore Alquier: From linear to non-linear broadcast contents: how to think the “audiovisual augmented archive”

Q&A

Track 2: Opera

European Union & GDPR

Chair: Renata Petrušić

Alejandra Michel: The legal framework for web archiving: focussing on GDPR

Tom Storrar & Chris Doyle: Creating an archive of European Union law for Brexit

Els Breedstraet: Setting up an EU web preservation service for the long-term – tales of a (sometimes) bumpy road

Marinos Papadopoulos, Charalampos Bratsas, Michalis Gerolimos, Konstantinos Vavoussis, Eliza Makridou & Dimitra

Hioti: Text and data mining for the national library of Greece in consideration of GDPR

Q&A



Check out abstracts!

#iipcWAC2019

15 YEARS HAW | Hrvatski arhiv weba
Croatian Web Archive

Organizers



Support



Venue

THE WESTIN
ZAGREB