

## Computational Archival Science

*Workshop Chairs: Mark Hedges, Victoria Lemieux, Richard Marciano*

Time	Title	Presenter/Author
9:00 – 9:15	Welcome	Mark Hedges (King's College London), Victoria Lemieux (U. British Columbia) Richard Marciano (U. Maryland),
9:15 – 10:00	Keynote	Dave de Roure
<b>10:00 – 10:40</b>	<b>SESSION 1: Analytics for archival processing</b>	
10:00 – 10:20	Automatic Extraction of Dublin Core Metadata from Presidential E-records	W. Underwood
10:20 – 10:40	Multi-label Classification of Chinese Judicial Documents based on BERT	M. Dai, C.L. Liu
<b>10:40 – 11:00</b>	<b>Coffee Break (coffee not supplied!)</b>	
<b>11:00 – 11:40</b>	<b>SESSION 2: Analytics for archival processing (cont.)</b>	
11:00 – 11:20	A Computational Approach to Historical Ontologies	M. Kelly, J. Greenberg, C. B. Rauch, S. Grabus, J. P. Boone, J.A. Kunze, P. Melville Logan
11:20 – 11:40	Digital Curation and Machine Learning Experimentation in Archives	T. Randby, R. Marciano
<b>11:40 – 12:40</b>	<b>SESSION 3: Analysing historical data and documents</b>	
11:40 – 12:00	Curation of Historical Arabic Handwritten Digit Datasets from Ottoman Population Registers: A Deep Transfer Learning Case Study	Y. Said Can, M. Erdem Kabadayi
12:00 – 12:20	HRCenterNet: An Anchorless Approach to Chinese Character Segmentation in Historical Documents	C-W Tang, C-L Liu, P-S Chiu
12:20 – 12:40	Towards Automatic Data Cleansing and Classification of Valid Historical Data: An Incremental Approach Based on MDD	E. O'Shea, R. Khan, C. Breathnach, T. Margaria
<b>12:40 – 1:40</b>	<b>Lunch Break (lunch not supplied!)</b>	
<b>1:40 – 2:00</b>	<b>Discussion (and questions for keynote speaker)</b>	
<b>2:00 – 2:40</b>	<b>SESSION 4: Representation in archives</b>	
2:00 – 2:20	Computational Treatments to Recover Erased Heritage: A Legacy of Slavery Case Study	L. Perine, R. K. Gnanasekaran, P. Nicholas, A. Hill, R. Marciano
2:20 – 2:40	Elevating "Everyday" Voices and People in Archives through the Application of Graph Database Technology	M. Conrad, L. Williams
<b>2:40 – 3:20</b>	<b>SESSION 5: Visual and audio archives</b>	
2:40 – 3:00	A Study of Spoken Audio Processing using Machine Learning for Libraries, Archives and Museums	W. Xu, M. Esteva, P. Cui, E. Castillo, K. Wang, H-R Hopkins, T. Clement, A. Choate, R. Huang
3:00 – 3:20	From Computational De-Morphogenesis to Contaminated Representation for the Contemporary Digital Tectonics and Lexicon: Vulnerability and Resilience of the Generative Representation in the Web-Oriented Platform – BIM/GD	A. De Masi
<b>3:20 – 3:40</b>	<b>Coffee Break (coffee not supplied!)</b>	
<b>3:40 – 4:20</b>	<b>SESSION 6: Web and social media archives</b>	
3:40 – 4:00	Modeling Updates of Scholarly Webpages Using Archived Data	Y. Jayawardana, A. C. Nwala, G. Jayawardana, J. Wu, S. Jayarathna, M.L. Nelson, C. Lee Giles
4:00 – 4:20	Using a Three-step Social Media Similarity (TSMS) Mapping Method to Analyze Controversial Speech Relating to COVID-19 in Twitter Collections	Z. Yin, L. Fan, H. Yu, A. Gilliland
<b>4:20 – 5:00</b>	<b>SESSION 7: Discussion</b>	
<b>5:00 – 5:20</b>	<b>Closing Remarks (Organizers)</b>	