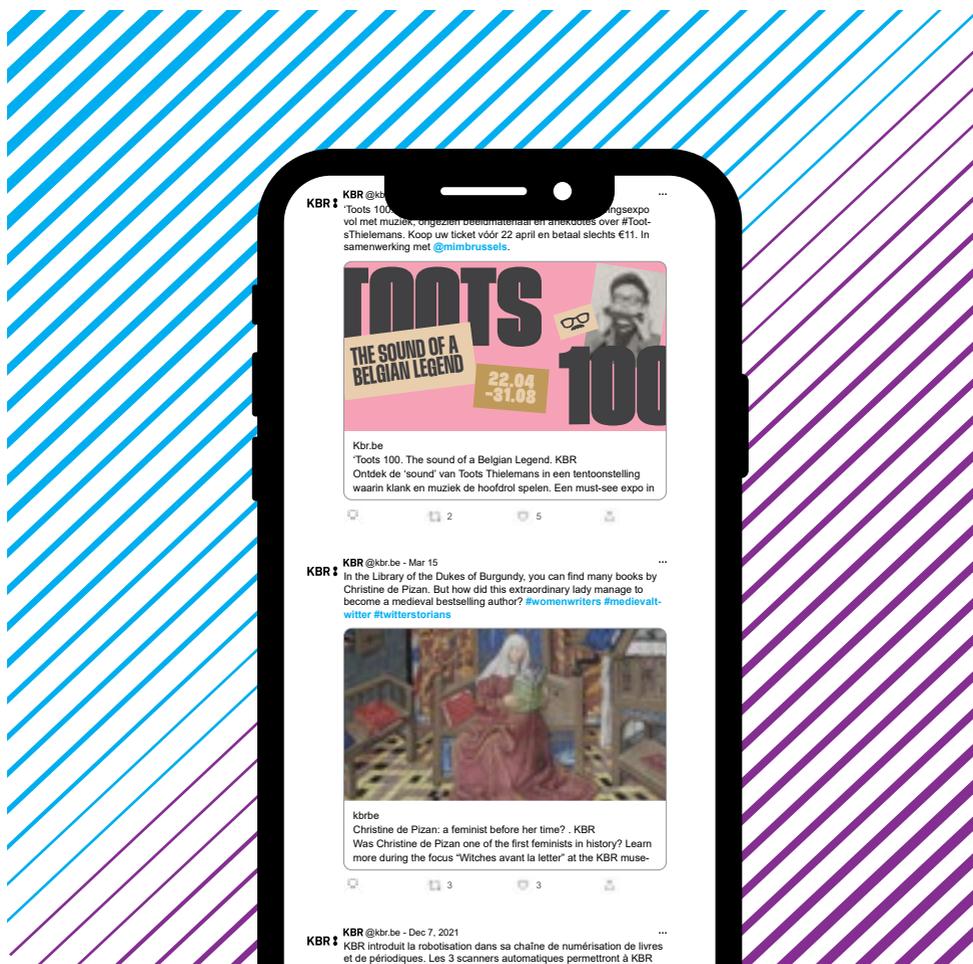


Wanted: Social Media Data

Archiving Practices and Research Use

Colloquium - KBR - 15.09.2022



This colloquium is organised with the support of the BRAIN-be programme at BELSPO and the Fonds de la Recherche Scientifique.

INTRODUCTION

Social media constitutes an integral part of our daily lives and even our culture. The content published on social media platforms holds a lot of potential and represents a non-negligible source of information for many researchers in the fields of history, political science, linguistics, sociology, digital humanities and information and computer science. In order to safeguard this content from disappearing and to open it up for research, many heritage institutions have started social media archiving programmes. Archiving social media data is a complex and challenging process from the point of view of capturing, preserving and making the collections accessible since the content is especially ephemeral. In addition there are a number of legal aspects to consider in this process.

To tackle this challenge, KBR - the Royal Library of Belgium, launched the BESOCIAL project in 2020 in collaboration with the universities of Louvain (Centre for Natural Language Processing), Namur (CRIDS - Research Centre in Information, Law and Society) and Ghent (Ghent Centre for Digital Humanities, IDLab - Internet Technology and Data Science Lab and mict - Research Group for Media, Innovation and Communication Technologies). This two-year research project funded by the BRAIN.be programme of the Belgian Science Policy Office (BELSPO) works to develop a sustainable strategy for archiving and preserving the Belgian social media sphere. The interdisciplinarity of the research team is one of the strengths of the project since technical, legal and operational aspects are covered as well as aspects related to user needs.

In an effort to summarise some of the findings of this project and bring experts together we will host the colloquium 'Wanted: Social Media Data - Archiving Practices and Research Use'. The potential of social media data for research and social media archiving practices will be discussed during the keynote and two panel discussions. International and Belgian speakers will share their expertise on topics such as using tweets as historical sources, lowering barriers to access the archived web, the opportunities and challenges of doing social media research and getting started with social media archiving with limited resources.

The colloquium is a time to come together and reflect further about social media archiving practices and research use. We would appreciate it if you can help us keep track of the conversations online by using the hashtag #BELsocialmedia.

WANTED: SOCIAL MEDIA DATA – ARCHIVING PRACTICES AND RESEARCH USE

PROGRAMME | 15.09.2022 | KBR

Morning

- 9.00 - 9.15: **Registration**
- 9.15 - 9.30: **Welcome**
Sara LAMMENS - KBR
- 9.30 - 10.15: **API or Archive? Tormented Ways to Transform Tweets into Historical Sources**
Keynote Prof. Dr. Frédéric CLAVERT – Université de Luxembourg
- 10.15 - 10.30: Coffee break
- 10.30 - 12.00: **BESOCIAL: Towards a Sustainable Strategy for Social Media Archiving at KBR**
Lise-Anne DENIS - UNamur, Pieter HEYVAERT - UGhent,
Fien MESSENS - KBR & UGhent, Eva ROLIN - UCLouvain,
Sophie VANDEPONTSEELE - KBR, Eveline VLASSENROOT - UGhent
- 12.00 - 13.30: Lunch break (buffet served in the Galerie room)
Separate lunch meeting with the members of the follow-up committee of the BESOCIAL project in the Consilium room

Afternoon

- 13.30 - 15.00: **Social Media Archiving from an Institutional Point of View (panel)**
Moderator: Asst. Prof. Dr. Julie M. BIRKHOLZ - KBR & UGhent
- Social Media Archiving at the Royal Danish Library - Reflections and Status Quo**
Anders KLINDT MYRVOLL – Royal Danish Library
- Towards Best Practices for the Archiving of Social Media by Private Archival Institutions in Flanders**
Katrien WEYNS – KADOC-KU Leuven
- Archiving the Web in Small Archive Centers, from Archives de Quarantaine to Mémoire de Confinement**
Virginien HORGE – Archives of the City of Mons

- 15.00 - 15:15: **Next Steps for Social Media Archiving at the IIPC**
Dr. Olga HOLOWNIA - International Internet Preservation Consortium (IIPC)
- 15.15 - 15.30: Coffee break
- 15.30 -17.00: **Research Use of Social Media (panel)**
Moderator: Dra. Eveline VLASSENROOT - UGhent
- Lowering Barriers to Accessing the Archived Web: The Archives Unleashed Project**
Prof. Dr. Ian MILLIGAN – Waterloo University
- Research Implications of Online Interaction: Research in Linguistics & Communication for Community Action**
Dr. Louise-Amélie COUGNON – Université Catholique de Louvain
- The Opportunities and Challenges of Social Media Research: Insights from the PERCEPTIONS Project**
Dr. Jamie MAHONEY – Northumbria University
- 17.00 - 17:15: **Closing Remarks**
Prof. Dr. Julie M. Birkholz – KBR & Ghent University

SPEAKERS

Asst. Prof. Frédéric CLAVERT

Trained in political sciences and international history, Frédéric Clavert wrote his PhD thesis on Hjalmar Schacht – German Minister of Economics (1934–1937) and President of the Reichsbank (1924–1930; 1933–1939). From one position in Luxembourg, to other ones in Paris and Lausanne, he gradually turned his attention to the study of the relationship between historians and their primary sources in the digital age (The allure of the archive in the digital era, a project led with Caroline Muller, Rennes 2) on the one hand, and the use of massive data from web platforms in memory studies on the other. He is today an Assistant Professor in contemporary history at the C2DH (University of Luxembourg) and managing editor of the Journal of Digital History.

Dr. Louise-Amélie COUGNON

Louise-Amélie Cougnon completed a PhD in sociolinguistics of electronic messages at the Cental (UCLouvain), a research center dedicated to natural language processing. After a few years of lecturing for several university and senior university linguistic courses, she became a researcher in Social and Political Sciences and studied online public debate for pedagogical, political and media purposes (Webdeb project). After 2 years, she started an FNRS project dedicated to the links between Social Media, Youth & New Skills. Now she leads the Research team in the MiIL, a UCLouvain lab dedicated to Media Innovation and Intelligibility, where she conducts an important project of Digital Corpora Monitoring.

Dr. Olga HOLOWNIA

Olga Holownia is the IIPC Senior Program Officer, based at the Council on Library and Information Resources. Olga manages the communications and programmes of the International Internet Preservation Consortium (IIPC). Her key projects include the organisation of the annual IIPC General Assembly, Web Archiving Conference, training events and webinars. She has been responsible for research engagement and outreach activities and co-chairs the IIPC Research Working Group.

Virginien HORGE

Virginien Horge is the Archivist of the City of Mons and member of the Association des Archivistes Francophone de Belgique (AAFB). He worked on the national «Archives de Quarantaine» project (AAFB) and on the «Mémoire de Confinement» project (City of Mons). One of the goals of these projects was to archive the web and social networks related to the pandemic.

Sara LAMMENS

Sara Lammens studied Musicology and Philosophical Academy at the Katholieke Universiteit Leuven. She started her professional career as a researcher at the Alamire Foundation, the international centre for the study of music in the Low Countries. In 2004 Sara began working for the Royal Library of Belgium (KBR). She started as Head of the Music Department, then switched to the Communications Department and in 2013 she became the Support Services Director of KBR. Since May 2017 she is the General Director a.i. of KBR, responsible for the general management of the institution and its representation in Belgium and abroad.

Anders KLINDT MYRVOLL

Anders Klindt Myrvoll is the Programme Manager at the national Danish web archive, [Netarkivet](#), at the [Royal Danish Library](#). Together with colleagues, he is collecting, preserving and providing access to the Danish web. Prior to web archiving Anders worked more than 13 years in the film and media industry, collaborating globally on high-end localisation, making original content for children, saving digital cultural heritage and much more. You can find him on [LinkedIn](#) or @andersklindt on [Twitter](#).

Dr. Jamie MAHONEY

Jamie Mahoney is a Research Fellow in the Northumbria Social Computing (NorSC) research group, in the Department of Computer and Information Sciences at Northumbria University, UK. His research interests include the use of social media in various contexts, and the ethical considerations and implications of using this data for research purposes. Jamie is currently involved in the PERCEPTIONS project, which is funded through the EU's Horizon 2020 programme.

Prof. Ian MILLIGAN

Ian Milligan is a professor of history at the University of Waterloo, where he is also Associate Vice-President, Research Oversight and Analysis. Milligan's primary research focus is on how historians can use web archives, as well as the impact of digital sources on historical practice more generally. Since 2017, Milligan has been principal investigator of the Archives Unleashed project.

Katrien WEYNS

Katrien Weyns studied history at KU Leuven and deepened her knowledge of heritage management during a master in archival science. As a digital heritage consultant in KADOC-KU Leuven she advises private organisations and individuals on the preservation of their digital archives and publications. Since 2019 she is head of the heritage library and manages the acquisition, preservation and disclosure of books and serials that were used by, published by or about religiously inspired persons and organisations. She developed and coordinates the Flemish heritage project 'Best practices for the archiving of social media in Flanders and Brussels (2020-2023)' that is conducted in partnership with meemoo and other Flemish heritage institutions. This project aims at researching the technical, legal and organisational framework of social media archiving in Flemish private archival and heritage institutions and developing sustainable solutions for these institutions to capture and archive social media.

BESOCIAL PROJECT TEAM

Asst. Prof. Dr. Julie M. BIRKHOLZ

Julie M. Birkholz is the Lead of the KBR's Digital Research Lab and Assistant Professor Digital Humanities at the Ghent Centre for Digital Humanities, Department of History, Ghent University. She is also the coordinator of the ESFRI - KBR Virtual Lab (2022 - 2025) that is developing a data-on-demand service and research infrastructure for KBR. She brings together expertise from the digital humanities working to facilitate access to the data level of the collections, and text and data mining research. Her research expertise is developing research workflows for extracting information from textual sources (i.g. text and data mining) and specifically for investigating (historical) social networks.

Dra. Sally CHAMBERS

Sally Chambers is Digital Humanities Research Coordinator at the [Ghent Centre for Digital Humanities](#) and coordinator of the [DATA-KBR-BE](#) project at [KBR](#) to facilitate data-level access to KBR's digitised and born-digital collections for digital humanities research. Sally has participated in the [PROMISE project](#) to develop a federal strategy for the establishment of a web-archive for Belgium, the [BESOCIAL](#) project related to social media archiving in Belgium and is also a member of the [WARCNet](#) research network.

Lise-Anne DENIS

After completing a Master's degree in law followed by an LL.M. in ICT law, Lise-Anne Denis is now working as a researcher at the Research Centre in Information, Law and Society (CRIDS), based at the University of Namur. She specialises in eCommerce, Media law and Privacy law. Within the BESOCIAL project, she participated in the analysis of all legal questions related to social media archiving.

Prof. Dr. Cécile DE TERWANGNE

Prof. Cécile de Terwangne is full professor at the Faculty of Law of the University of Namur. She teaches courses in Computer and Human Rights, Data Protection and Cybercrime. She is director of the post-graduate Program in IT Law. She is research director at the CRIDS (Research Centre for Information, Law and Society). She took part in numerous European and national research projects in the fields of data protection, privacy, freedom of information, e-government, re-use of public sector information, etc. She is an appointed expert for the Council of Europe and the European Commission. She is member of the Reference Panel of the Global Privacy Assembly.

Friedel GEERAERT

Friedel Geeraert is an expert in web archiving at KBR. She previously worked as a researcher on the [PROMISE project](#) that focused on the development of a Belgian web archive at the federal level. She is also involved in the [BESOCIAL](#) project as coordinator, is a member of the [WARCNet](#) research network and represents KBR at the [IIPC](#) (International Internet Preservation Consortium). She holds a Master's degree in History (KU Leuven) and in Information and Communication Science and Technology (Université Libre de Bruxelles).

Dr. Pieter HEYVAERT

Pieter Heyvaert is a development lead, developer advocate, and diversity & inclusion ambassador at Ghent University - IDLab, imec. He obtained the degree of Doctor of Computer Science Engineering in 2019 from Ghent University. His interests are in the scope of the Semantic Web. More specific, his PhD research focused on improving the effectiveness of the creation and execution of knowledge graph generation rules.

Fien MESSENS

Fien Messens is an expert in social media archiving at KBR. She is project leader of the BESOCIAL research project, which aims to develop a sustainable strategy for archiving and preserving social media at KBR. She holds a Master's degree in Art History and Digital Humanities and is currently applying her degree to KBR's born-digital collections in the framework of the BESOCIAL project.

Dr. Peter MECHANT

Peter Mechant holds a PhD in Communication Sciences from Ghent University (2012) and has been mainly working on a variety of research projects related to e-gov (open and linked data), smart cities, social software and online communities. As senior researcher, he is involved in managing projects and project proposals on a European, national as well as regional level.

Dra. Alejandra MICHEL

Alejandra Michel is a senior researcher at the Research Centre in Information, Law and Society of the University of Namur, in Belgium (CRIDS/NaDI, UNamur). She is also Head of the "Media unit" dedicated to research on various aspects related to the gathering, production, dissemination, moderation and protection of content in the digital and audio-visual worlds. Alejandra also teaches media law, fundamental rights law and digital aspects of privacy law such as video surveillance (CCTV cameras) and social networks. Since March 2019, she is also a Member of the "Conseil de Déontologie Journalistique" (press council in French-Speaking Belgium) for the civil society.

Eva ROLIN

After a bachelor's degree in Romance languages and literature at the University of Namur, Eva Rolin decided to do a master's degree in automatic language processing at the University of Louvain. She then had the chance to join the Cental (Centre du Traitement Automatique du Langue) to work on various projects: writing articles on automatic lexical and syntactic simplification, participating in various projects such as Aidalex and of course BESOCIAL in which she puts her linguistic and computer skills to use.

Sophie VANDEPONTSEELE

Sophie Vandepontseele has a degree in History. She has worked for many years as an archivist and records manager. She currently works as Director of Contemporary Collections at KBR and is, among other things, the promoter of several research projects, including BESOCIAL, all of which have the ultimate goal of creating new digital collections and making them as accessible as possible to a wide audience.

Dra. Eveline VLASSENROOT

Eveline Vlassenroot joined the research group for Media, Innovation and Communication Technologies (imec-mict-UGent) after obtaining a Master's degree in Communication Sciences and a Bachelor's degree in Information Management in 2017. She is, among other things, involved in projects related to (the use of) web archives. Furthermore she is active in the Flemish Interoperability Program OSLO (Open Standards for Linking Organisations) which focuses on building bridges between information within and outside the government.

Dr. Patrick WATRIN

Patrick Watrin became operational manager of the CENTAL, after completing a PhD at UCLouvain, a postdoctorate at the Gaspard Monge Institute and setting up a spin-off, Earlytracks, active in the medical information sector. For almost 20 years, he has conducted research in the field of information extraction and structuring and is now interested in the study of neural networks and deep learning to increase the efficiency of analysis tools on noisy (real) data.

NOTES

