



HAL
open science

Exploring Polyvocality in Online Narratives of French Colonial Memory through Web Archives: Methodological and Ethical considerations

Mara Bertelsen, Sophie Gebeil, Véronique Ginouvès, Christine Mussard

► **To cite this version:**

Mara Bertelsen, Sophie Gebeil, Véronique Ginouvès, Christine Mussard. Exploring Polyvocality in Online Narratives of French Colonial Memory through Web Archives: Methodological and Ethical considerations. Web archives in context, Apr 2024, Paris, France. halshs-04580170

HAL Id: halshs-04580170

<https://shs.hal.science/halshs-04580170v1>

Submitted on 19 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open licence - etalab

EXPLORING POLYVOCALITY IN ONLINE NARRATIVES OF FRENCH COLONIAL MEMORY THROUGH WEB ARCHIVES

Methodological and ethical considerations

A European network on the Decolonization of Digital Audio Visual Archives

Polyvocal Interpretations of Contested Colonial Heritage (PICCH) is a European project involving five national partners and coordinated by Prof. Daniela Petrelli (Sheffield Hallam University). Alexander Badenoch (Utrecht University) is responsible for the Dutch team, while Pia Borlund (Oslo Metropolitan University) leads the Swedish team. PICCH explores how archival material created in a colonial mindset can be reappropriated and reinterpreted to become an effective source for decolonization and the basis for a future inclusive society.

A transmedia and interdisciplinary approach

The French team at Aix-Marseille University, coordinated by Sophie Gebeil (PI), comprising Véronique Ginouvès (archivist), Christine Mussard (historian), Pauline Savéant (Ph.D. student) and Mara Bertelsen (research assistant), in partnership with the INA (Institut National de l'Audiovisuel), mobilizes diverse archives to study colonial and postcolonial narratives. French web archives are crucial for studying the evolution of the multiple narratives published online in recent years on the polyvocal memories of colonization in the Maghreb and its postcolonial consequences.

Focus: The online remembrances of the 1983 March for Equality and Against Racism

Based on a corpus composed of data obtained from the web archive of the INA, which includes web pages referencing the history of the 1983 March for Equality and Against Racism, we explored how this postcolonial event was portrayed online, improving computational methods for text analysis. The study of large corpora of historical data from archived web sources is grounded in a corpus co-constructed with web archivists and focused on the year of the March's thirtieth anniversary in 2013. Semantic network analysis underscores the different interpretations of this historical event.



“Retrospective on the March Against Racism leaving from Marseille and arriving in Paris”, Broadcast Archive, FR3, Ina, 1983-12-03. Id: RAC240-0448

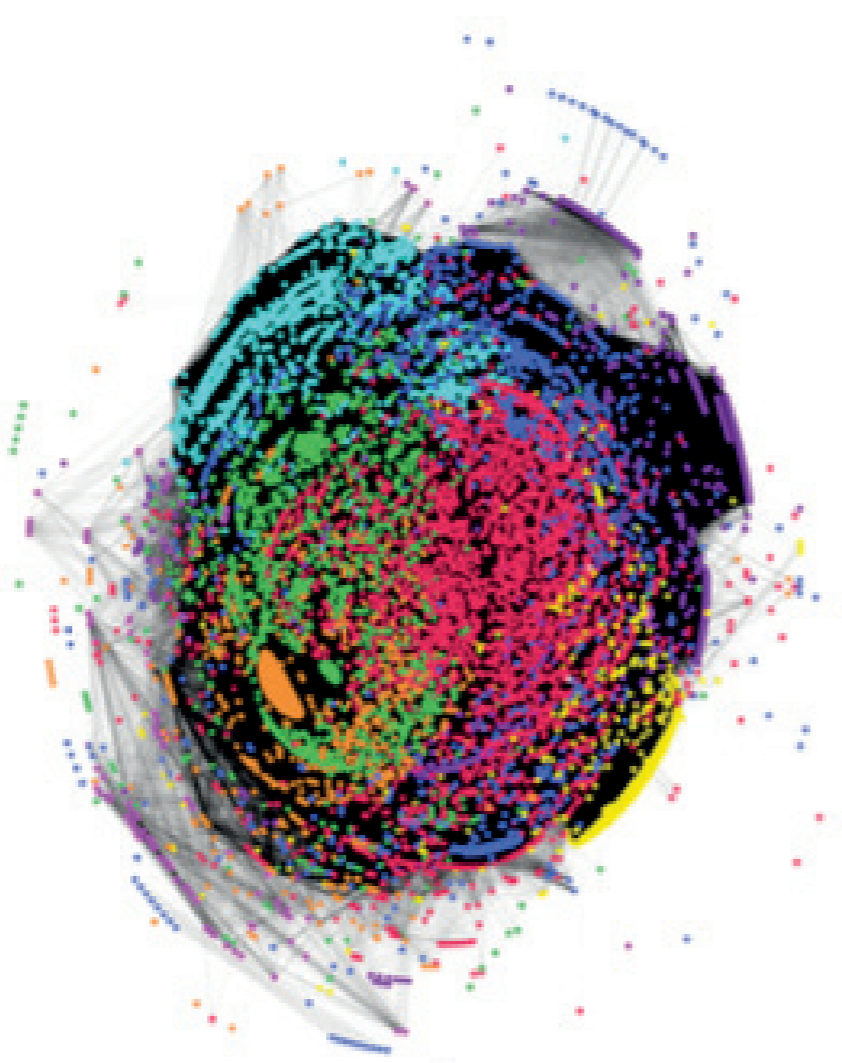
How does it work?

```
«content»: «ee601d865509c75dc911378d6e9fcf1fd5bbf0a1cdd466bc3e3ded306da6b153»,
«client_country»: «fr»,
«client_ip»: «10.3.209.54»,
«page»: «1»,
«inline_level»: «0»,
«status»: «ok»
```

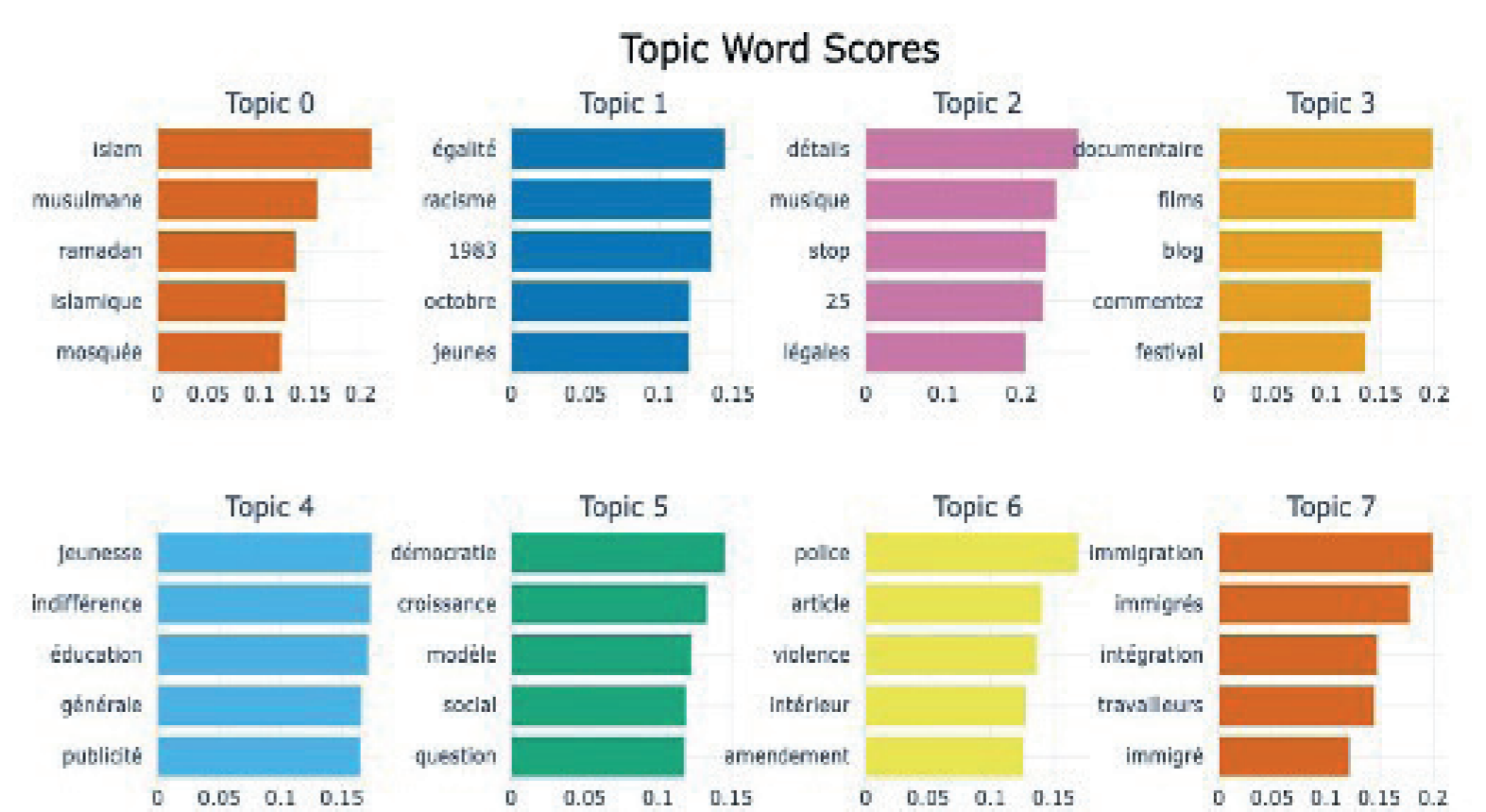
An element of a website in json format, with indentation and line breaks to make it easier to read.

Do NOT CHANGE		to validate	
entity	original (dictionary)	entity	normalized (dictionary)
.	. I. A. Pentago (LOC)	.	. I. A. Pentago
'1	'1 (LOC)	'1	'1
'10	'10 (LOC)	'10	'10
'10e	'10e arrondissement de Paris (LOC)	'10e	'10e arrondissement de Paris
'10ème	'10ème arrondissement de Paris (LOC)	'10ème	'10ème arrondissement de Paris
'11	'11 (LOC)	'11	'11
'11e	'11e arrondissement de Paris (LOC)	'11e	'11e arrondissement de Paris
'11ème	'11ème arrondissement de Paris (LOC)	'11ème	'11ème arrondissement de Paris
'11ème	'11ème arrondissement de Paris (LOC)	'11ème	'11ème arrondissement de Paris
'11ème	'11ème circonscription des Français de l'Etranger (LOC)	'11ème	'11ème circonscription des Français de l'Etranger
'12e	'12e arrondissement de Paris (LOC)	'12e	'12e arrondissement de Paris
'13	'13 (LOC)	'13	'13
'13	'13 arrondissement de Paris (LOC)	'13	'13 arrondissement de Paris
'13 Nord	'13 Nord (LOC)	'13 Nord	'13 Nord
'13e	'13e arrondissement de Paris (LOC)	'13e	'13e arrondissement de Paris
'13e	'13e circonscription de Paris (LOC)	'13e	'13e circonscription de Paris
'13e	'13e circonscription des Hauts de Seine (LOC)	'13e	'13e circonscription des Hauts de Seine

Extracting code by machine is not sufficient. Human intervention is also necessary to clean the data in accordance with scientific objectives, and this cleaning process is documented as well. The results are still undergoing analysis.



Document-Document (entities) Network's Louvain communities (Davide Rendina)



Topic Modelling. Graph created by Davide Rendina (dir. S. Gebeil, M. Géniois, P. Bellot).

Conclusion

Beyond the results, the project demonstrates the strength of multidisciplinary work and collective efforts. The analysis of a corpus extracted from web archives is slated for publication in the proceedings of the RESAW conference Exploring the Archived Web During a Highly Transformative Age (eds. S. Gebeil and J.-C. Peyssard), Firenze University Press, forthcoming end 2024.

Poster created by Maja Mellet TELEMMe, PICCH, 26 April 2024, IIPC, Paris, France. Poster under Etalab licence.



Maison méditerranéenne des sciences de l'homme UAR 3125

