



IIPC WEB ARCHIVING CONFERENCE 2021

Brazilian elections web archive

Moisés Rockembach

Federal University of Rio Grande do Sul

Research Group on Web Archiving and Digital Preservation

moises.rockembach@ufrgs.br



The Research Group of Web Archiving and Digital Preservation was created in 2017 to investigate the characteristics of web archiving through national and international initiatives, dealing with both the policies and the technologies involved in the process.

Research Group on web archiving and digital preservation in Brazil

www.ufrgs.br/nuaweb

Brazilian Web Archive

www.arquivo.org.br



Research Group on web archiving and digital preservation in Brazil

www.ufrgs.br/nuaweb

Brazilian Web Archive

www.arquivo.org.br

Dissemination: lectures, podcasts..



Rodrigues, V. L. D. .; Rockembach, M. Web archives as a historiographical source. RDBCI: Digital Journal of Library and Information Science, 2021

MELO, Jonas Ferrigolo, and ROCKEMBACH, Moisés. Archivability of websites for digital preservation: study from the health area, Revista Eletrônica de Comunicação, Informação e Inovação em Saúde

Martins, M. R., & Rockembach, M. (2020). Mapping of audiences for academic web archiving initiatives. Intercom: Revista Brasileira de Ciências da Comunicação, 43(1), 71-88.

Martins, M. R.; Rockembach, M. Promoção de iniciativas de arquivamento da web: um estudo a partir da rede de públicos estratégicos da UFRGS. In: II Congresso de Gestão Estratégica da Informação, Empreendedorismo e Inovação. Porto Alegre, 2019

Ferreira, L. B. ; Martins, M. R. ; ROCKEMBACH, M.. Arquivamento da web e seus usos na comunicação científica. PRISMA.COM, v. 36, p. 1, 2018.

Rockembach, M. Arquivamento da Web: estudos de caso internacionais e o caso brasileiro. Revista Digital de Biblioteconomia e Ciência da Informação, 2018.

Rockembach, M.. Web Archiving: international case studies and the Brazilian case (English Version) Digital Journal of Library and Information Science, v. 16, p. 1, 2018

Rockembach, M. A web como memória: preservação digital e arquivamento da web no Brasil e no exterior. I Congresso Internacional em Humanidades Digitais, Rio de Janeiro, 2018

Rockembach, M.; Pavao, C. M. G. . Políticas e tecnologias de preservação digital no arquivamento da web. Rev. Ibero Americana de CI, 2018.

Rockembach, M.. Inequalities in digital memory: ethical and geographical aspects of web archiving. International Review of Information Ethics, v. 26, p. 1, 2017.

Research Group on web archiving and digital preservation in Brazil

www.ufrgs.br/nuaweb

Brazilian Web Archive

www.arquivo.org.br

Some publications..

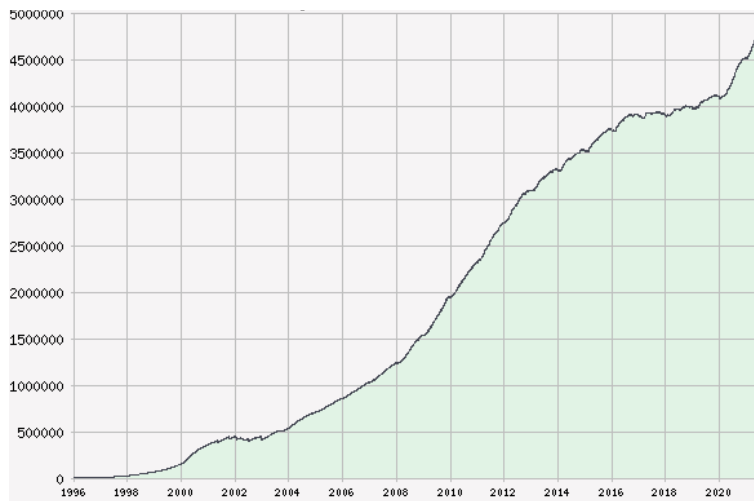
Background

- Infrastructure for Digital Memory and Digital Humanities
- Web as the archives of contemporary life (Brügger and Finnemann, 2013)
- Patterns of large-scale political action (Weber 2018)
- Importance of establishing web archiving policies in the formation of a digital web memory in Brazil (Rockembach and Pavão, 2018)

Context - origins of the web in Brazil

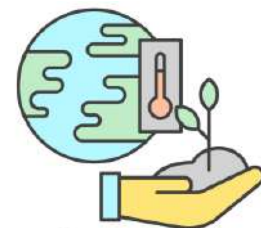
- ❑ National Education and Research Network (*Rede Nacional de Ensino e Pesquisa* - RNP) – founded in 1989
- ❑ Ordinance number 13, of April 20, 1995 (Brazil 1995) - Regulated the use of the public telecommunications network for internet connection services
- ❑ Creation of the Internet Management Committee in Brazil (Comitê Gestor da Internet no Brasil) - Interministerial Ordinance number 147, of May 31, 1995

Increase of .br domains



.br domains - 1996 - 2021

- Our scope:
 - ▣ Elections
 - ▣ Institutional
 - ▣ Climate changes
 - ▣ Covid-19 (since april 2020)



The 2018 Brazilian presidential elections and the preservation of web campaigns

- Several organizations and countries are preserving elections on the web (USA, Australia, Portugal, Chile...
- LoC - web archiving of elections: Philippines, Timor-Leste and Indonesia. Brazil (2010)
- Collaborative efforts are needed to extend this coverage
- Our research: Brazilian presidential elections of 2018

BRAZILIAN ELECTIONS 2018



Main goal: to analyze the production of information in the web environment and the respective preservation of digital memory about the **Brazilian presidential elections of 2018**.

Specific goals: identify the bibliographic references, digital curation, collect and store websites, based on web archiving standards, and analyze the web archive content.

Methods

Four phases:

1. Literature review

web archiving, digital preservation, digital curation, elections

2. Infrastructure configuration and software tests

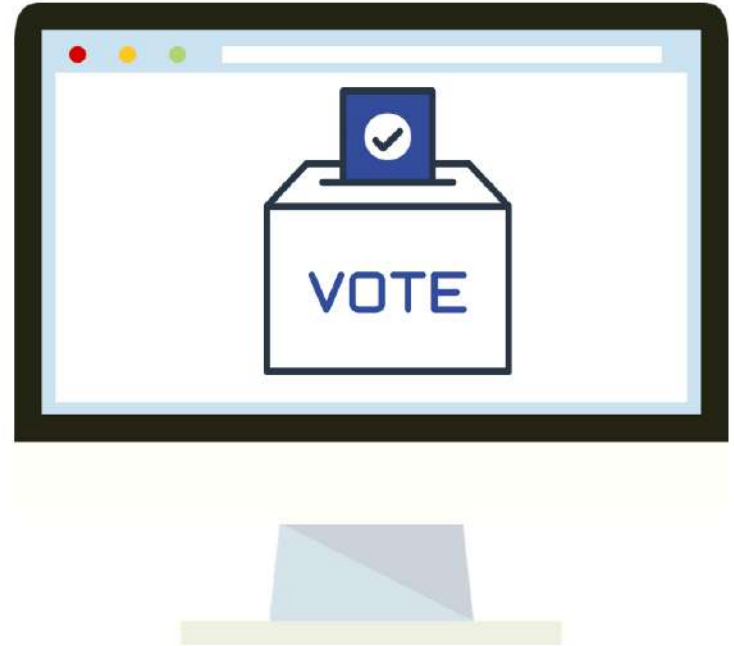
3. Data collection

4. Analysis of the collected data

Political Party	Seeds (candidate and political party)		
	Main website	Secondary website	Party website
12 - PDT	X	X	X
13 - PT	X	X	X
15 - MDB	X	-	X
16 - PSTU	-	-	X
17 - PSL	X	X	X
18 - REDE	X	-	X
19 - Podemos	X	-	X
27 – Democracia Cristă	X	-	X
30 - Novo	X	-	X
45 - PSDB	X	-	X
50 - PSOL	X	-	X
51 - Patriota	X	-	X
54 - PPL	-	-	X

Curation results

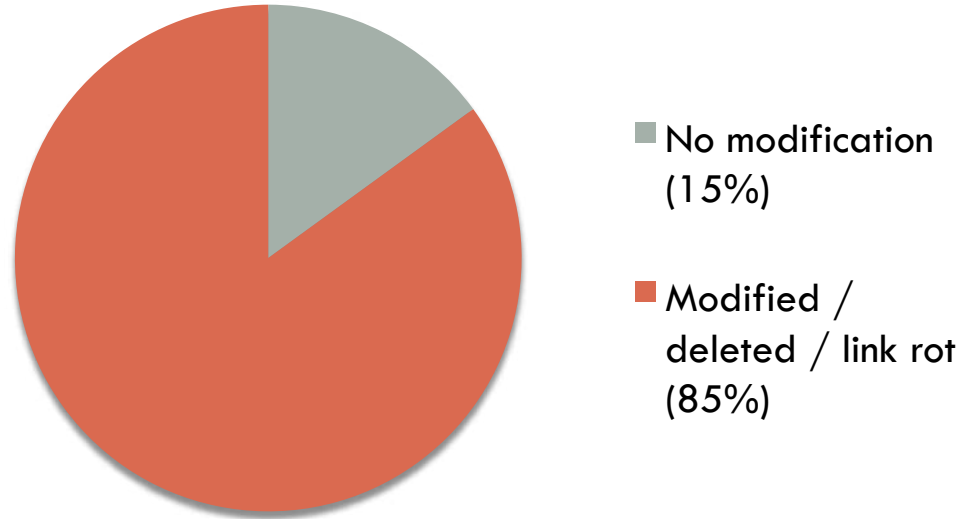
- 27 seeds, composed of 13 political party websites, 11 presidential candidate websites and 3 candidate special websites



Quality assurance

- Crawler problems - around 15% of captured webpages
- Quality assurance, focus on the content (Hockx-Yu, 2012)
- When possible, adopt standards that facilitate the capture of content. Archivability (Kelly et. al 2013; Banos and Manolopoulos 2015)

Changes over time (2018-2020)



Changes over time (2018-2020)



Modified



Não é possível acessar esse site

Não foi possível encontrar o endereço IP do servidor de www.geraldaalcmin.com.br.

DNS_PROBE_FAILED_NXD001AN

Offline

Recarregar



Error 5xx

Error 522
Connection timed out



What can I do?
If you're a visitor of this website:
Please try again in a few minutes.
If you're the owner of this website:



TUDO POSICIONAMENTOS APOIOS COMBATE A FAKE NEWS



Self-Archived



TUDO POSICIONAMENTOS APOIOS COMBATE A FAKE NEWS



arquivo.org.br – The Brazilian Web Archive

ARQUIVO DA WEB BRASILEIRA

Busque sites do passado

[Página Inicial](#)

[Navegue na web do passado](#)

[Arquivo da Web | Blog](#)

[Contribua](#)

[Sobre nós](#)



Pesquise

Utilize as buscas disponíveis no Arquivo da Web Brasileira, novas funcionalidades são implementadas constantemente.



Acesse

Nossas bases de dados estão disponíveis publicamente e de forma gratuita, para usos não comerciais, com objetivos de pesquisa, educacionais e científicos.



Contribua

A partir de sugestão de sites a serem preservados ou como voluntário para a melhoria do serviço do Arquivo da Web Brasileira.

- Make web archives available for search & research;
- Engage users and researchers community, namely in digital humanities;
- Education and training in digital preservation related to web archives.

Conclusion and future work

- The relevance of digital preservation and the future use by researchers and citizens / The urgency in digital preservation
- Elections 2022
- Engage users
- Expand the scope of the study in the web archives

References

BANOS, V., MANOLOPOULOS, Y. Web Content Management Systems Archivability. In East European Conference on Advances in Databases and Information Systems (pp. 198-212). Springer, Cham, 2015

BANOS, V., MANOLOPOULOS, Y. A quantitative approach to evaluate Website Archivability using the CLEAR+ method. International Journal on Digital Libraries, 17(2), 119-141, 2016

BERNERS-LEE, Tim. Information management: a proposal. Switzerland: CERN, 1989. Retrieved: <http://cds.cern.ch/record/369245/files/dd-89-001.pdf>

BRASIL. Portaria no 13, de 20 de abril de 1995. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 24 Abr. 1995a. Seção 1, p. 5666.

BRASIL. Portaria Interministerial no 147, de 31 de maio de 1995. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 31 maio 1995b

BRÜGGER, Niels; FINNEMANN, Niels Ole. The web and digital humanities: theoretical and methodological concerns. Journal of Broadcasting and Electronic Media, v. 57, n.1, 2013

BRUNELLE, J. F., et.al . The impact of JavaScript on archivability. International Journal on Digital Libraries, v. 17, n. 2, p. 95-117, 2016.

COSTA, Miguel; GOMES, Daniel; SILVA, Mário. The evolution of web archiving. International Journal on Digital Libraries, p. 1-15, 2016

LIBRARY OF CONGRESS, Retrieved <https://www.loc.gov/programs/web-archiving/for-researchers/>

HOCKX-YU, Helen. How good is good enough? – quality assurance of harvested web resources, 2012. Retrieved: <http://blogs.bl.uk/webarchive/2012/10/how-good-is-good-enough-quality-assurance-of-harvested-web-resources.html>

KELLY, M.; et. al. On the change in archivability of websites over time, International Conference on Theory and Practice of Digital Libraries. Springer, Berlin, Heidelberg. 2013. p. 35-47

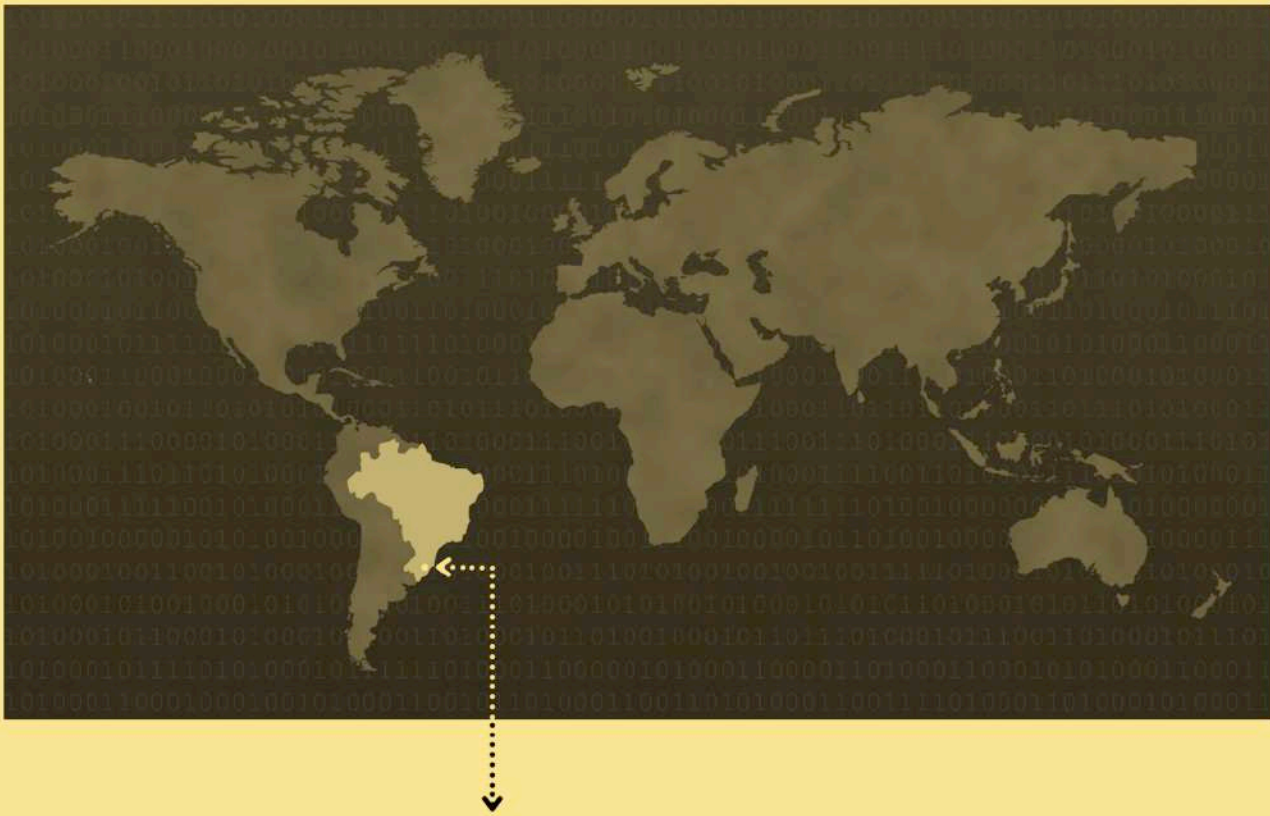
MILLIGAN, I., RUEST, N., LIN, J. Content selection and curation for web archiving: The gatekeepers vs. the masses. In Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries (pp. 107-110). ACM, 2016

NTOULAS, Alexandros, CHO, Junghoo, OLSTON, Christopher. Whats new on the web?: the evolution of the web from a search engine perspective. Proceedings of the 13th international conference on World Wide Web. ACM, 2004

ROCKEMBACH, M.. Web Archiving: international case studies and the Brazilian case. Digital Journal of Library and Information Science, v. 16, p. 1, 2018

WEBER, M. S.. Methods and approaches to using web archives in computational communication research. Communication Methods and Measures, 12(2-3), 200-215, 2018

Thank you!



Research Group on web archiving and digital preservation in Brazil

www.ufrgs.br/nuaweb

Brazilian Web Archive

www.arquivo.org.br

Moisés Rockembach

moises.rockembach@ufrgs.br

[@mrockembach](https://twitter.com/mrockembach)