

Technological study of Brazilian government websites

Newton Calegari / Reinaldo Ferraz – CEWEB.br





Newton Calegari



Reinaldo Ferraz

cewebbr nicbr cgibr

cewebbr nicbr cgibr

սորող

Introduction

cewebbr nicbr cgibr

• Colect webpages under .gov.br Top Level Domain

սորող

• Perform validation

ഞ

Technical enviroment

cewebbr nicbr cgibr

 WIRE-Nic is a fork of the open souce project WIRE (Web Information Retrieval Environment) that was developed by Center for Web Research from University of Chile. This fork brings some bug fixes and modifications to the original system.

mmmm

- A web crawler.
- Tools for extracting statistics from the collection.
- Tools for generating reports about the collection.

https://sourceforge.net/projects/wire-nic/

Technical enviroment

cewebbr nicbr cgibr

• ConNeCTOR - Convenient Network Characteristics Testing

Organized Routines

• Created to support the carrying out of tests related to geo-

localization of the servers and adherence to HTML and

mmm

accessibility (eMAG / WCAG) standards.

https://sourceforge.net/projects/connector-nic/

Collecting Webpages

cewebbr nicbr cgibr

- Some conditions the website should satisfy:
 - owning a domain with identified TLD .gov.br;
 - domain identified with the letters –uf.gov.br, that belong to different states of the Brazilian federation.

http://www.prefeitura.sp.gov.br/ http://www.nic.br/ http://www.google.com.br/

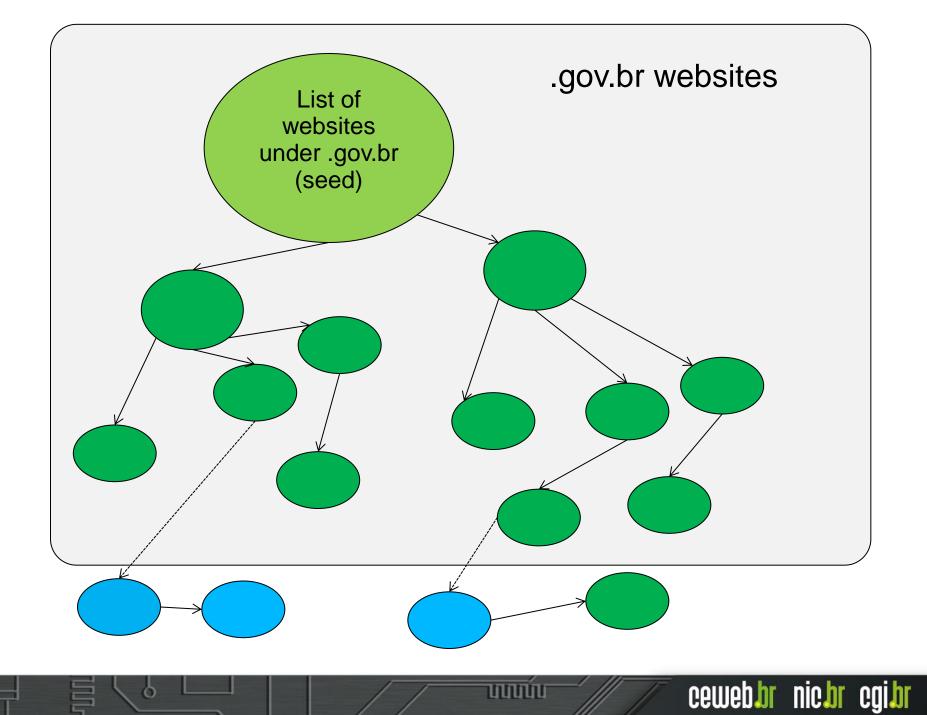
Innnn

Collecting Webpages

cewebbr nicbr cgibr

 The information in the domains is supplied by the authority for the register of domains in Brazil, Registro.br, with the authorization of the minister for planning, responsible for the use of the domains under .gov.br, and by the companies linked to the respective state governments.

սորող



Validation tools

- W3C Validator: Check W3C Standards compliance
 - <u>http://validator.w3.org</u>
- ASES Avaliador e Simulador de Acessibilidade em Sítios

mmm

<u>http://asesweb.governoeletronico.gov.br/ases/</u>

the Approace	(Dewel)	ngul
automatically)	12	T Only # moing
(especimental)		I' one message
Group Error Messages by Ty	pe	
n up blanup with in TML Tary		
Note error pages		T. Verbore Culpat
	expectmental) Group Error Mencages by Tr rup Intelloop with #TML Toty	espectmental)

A22-Controls Alexanity per http://doi.org	Intell MA
vano Litar Avenatoro Modalono furgenerati sicta	
	govbi
a)	(•) 94
Addentition Exercision and Addentition and Addentit	- (4000
gra	
Tomerar per espinateres incluints parts states income memory representações prótoas detento. Separatemente de pontes de properanção, e reseçuidores o institut apalitores, para sano velo	eclasies vandedes, OPv anavantes, insperie elititative reent
tion a fractional managine	
22	
24	
These Co.	
and in First .	
2014 Called The Second State of Second State o	
C. L. M. L. L. Partici Societies in Research in Research	
Alex Tele Levents' Arcone	A
	And the second s



Data analyzed

cewebbr nicbr cgibr

mmm

- Size of the Brazilian web
- Language of Brazilian web pages
- Geographical location of servers
- W3C standards compliance
- e-Mag standards compliance

• Size of the Brazilian web – in Gigabites

²⁰¹¹ 169 GB

õ

²⁰¹⁵ 286 GB

cewebbr nicbr cgibr

սսսսս

cewebbr nicbr cgibr

• Size of the Brazilian web – number of websites

²⁰¹¹ 2015 **12.891 11.080**

mmm

cewebbr nicbr cgibr

• Size of the Brazilian web – number of websites

The average size of websites in 2015:25 MBBiggest website in 2015:4.6 GB

Innnn

cewebbr nicbr cgibr

• Size of the Brazilian web – number of webpages

սորող

²⁰¹¹ 6.334.054

ഞ

²⁰¹⁵ 8.323.478

cewebbr nicbr cgibr

• Webpage language – Webpages in portuguese

2011 **97%**

2015 98.25%

սսսսս

cewebbr nicbr cgibr

• Number of servers located in Brazil

2011 **93%**

2015 **90%**

սսսսս

• Webpages which adhere to W3C standards

²⁰¹¹ 5,02%

Ó

²⁰¹⁵ 4,56%

սոողո



cewebbr nicbr cgibr

• Webpages which adhere to W3C standards

2015:

Pages with less than 10 erros:21%Pages with more than 100 error:10%

mmm

 Webpages which adhere to e-Mag standards (web accessibility)

²⁰¹¹

Ó

²⁰¹⁵5,7%

սոսող



Future work

cewebbr nicbr cgibr

Update the validator ASES to reach eMag 3 and WCAG
2.0

սորող

Check other HTML5 features

Thank you www.ceweb.br



June, 11th, 2015

www.nic.br | www.cgi.br