

Observatório da Acessibilidade da Web Brasileira

Rafael Pardini
João Bárbara
Henrique Scheid
Adriano C. M. Pereira
Wagner Meira Jr.
rafaelguedes@dcc.ufmg.br
jvictor@dcc.ufmg.br
henrique.furst@dcc.ufmg.br
adrianoc@dcc.ufmg.br
meira@dcc.ufmg.br

Departamento de Ciência da Computação
Universidade Federal de Minas Gerais
Belo Horizonte, MG, Brasil

Reinaldo Ferraz
Beatriz Rocha
reinaldo@nic.br
biarocha@nic.br

Centro de Estudos sobre Tecnologias Web - CEWEB.br
NIC.BR / CGL.br
São Paulo, SP, Brasil

ABSTRACT

The World Wide Web emerged more than three decades ago, and even after all this time, accessibility is still challenging and a major obstacle for it to be universal and inclusive. Despite W3C efforts to standardize and publish best practices towards a more accessible Web, we still lack methodologies and tools to collect Web content, measure and assess accessibility at large scale. In this work, we describe the Observatório da Acessibilidade Brasileira, a tool that supports accessibility assessment, aggregates global indicators and provides recommendations for improving accessibility on websites. The current version complies with eMAG accessibility standards and extends and integrates existing free software. We validate and demonstrate the effectiveness of the tool by assessing the accessibility of the Brazilian government websites. The tool roadmap includes adding new accessibility compliance analyses, improving data visualization and enriching website seed.

KEYWORDS

World Wide Web, Accessibility, Web crawler, ASES, WCAG

1 INTRODUÇÃO

A Web surgiu como uma grande plataforma de interface para a Internet. A proposta de Tim Berners-Lee de uma tecnologia que conectava documentos por meio de hiperlinks [10] mudou a forma como utilizamos serviços online. Em 1991, Tim Berners-Lee publicava a primeira página Web [3] e em 1994, lançava o *World Wide Web consortium*, uma entidade para desenvolver padrões para essa nova tecnologia chamada Web.

Três anos depois é lançada a Iniciativa de Acessibilidade na Web (WAI) com o objetivo de garantir que os padrões desenvolvidos pelo consórcio contemplem a acessibilidade na Web. No momento do seu lançamento, Tim Berners-Lee disse: "O W3C está empenhado em remover as barreiras de acessibilidade para todas as pessoas

com deficiência - incluindo pessoas surdas, cegas, com deficiência motora e cognitivas ou deficiências visuais. Planejamos trabalhar agressivamente com o governo, a indústria e os líderes comunitários para estabelecer e atingir as metas de acessibilidade da web" [11]. Dessa forma, desde sua criação, a acessibilidade é contemplada no desenvolvimento de padrões para a Web. Tornar a Web um ambiente sem barreiras de acesso para pessoas com deficiência sempre foi uma preocupação do W3C.

Atualmente existem diretrizes internacionais de acessibilidade na Web desenvolvidas no âmbito do W3C para a eliminação de barreiras de acesso para pessoas com deficiência. As *Web Content Accessibility Guidelines (WCAG)* [12] orientam gestores e desenvolvedores sobre como tornar o conteúdo acessível na Web. Esse é o principal documento de diretrizes técnicas para a acessibilidade na Web, que auxilia na formulação de políticas públicas e indica a conformidade de sites com políticas de acessibilidade. Diversos países utilizam o WCAG como padrão mínimo requerido em sites governamentais. Os Estados Unidos, por exemplo, exigem nível de conformidade baseado no WCAG pela *Section 508* [1].

No Brasil, os sites governamentais devem seguir as diretrizes do Modelo Brasileiro de Governo Eletrônico (eMAG) [7], que também foi desenvolvido com base nas orientações de acessibilidade do W3C. Para a verificação automática, existe a ferramenta ASES [5], desenvolvida pelo Governo Federal para testar páginas em conformidade com o eMAG. Também vale notar que a Lei Brasileira de Inclusão [9] é muito clara no capítulo 2, artigo 63, quando aponta a obrigatoriedade da acessibilidade em sites seguindo diretrizes internacionais de acessibilidade.

Mensurar e avaliar a acessibilidade não é uma tarefa simples, pois não se deve tratá-la de forma binária. Uma página sem erros de verificação pode ter sérios problemas de acessibilidade, da mesma forma que uma página com alguns erros pode não ter barreiras graves. Mesmo quando a conformidade com padrões de acessibilidade é satisfeita, só é possível atestar a acessibilidade de um site ou sistema depois da verificação humana, para detectar eventuais barreiras de acesso.

Mesmo com todos os esforços de padronização e legislação, é preciso verificar na prática se essas diretrizes estão sendo cumpridas, e o quão perto ou longe estamos do ideal. Faltam metodologias e

In: XX Workshop de Ferramentas e Aplicações (WFA 2021), Minas Gerais, Brasil. Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2021.

© 2021 SBC – Sociedade Brasileira de Computação.
ISSN 2596-1683

ferramentas que traçam panoramas gerais sobre a situação dos sítios alvos e que apontem os caminhos a serem seguidos, principalmente no cenário brasileiro. Para isso, o objetivo da ferramenta proposta e apresentada neste trabalho é coletar um grande volume de sítios da Web, medir indicadores de acessibilidade e apresentar tais resultados para os usuários e desenvolvedores. A versão atual da ferramenta apresenta o ASES como analisador da conformidade dos sítios, considerando a versão 3 do Modelo de Acessibilidade em Governo Eletrônico (eMAG).

2 TRABALHOS RELACIONADOS

2.1 Observatório Português da Acessibilidade

O Observatório Português da Acessibilidade [2] é um portal que visa verificar sites de institutos públicos, governamentais e não governamentais e apresentar estatísticas globais de acordo com diretrizes portuguesas e internacionais. Sua metodologia institui testes manuais de acessibilidade e usabilidade, incluindo cidadãos com necessidades especiais. Outro aspecto são as funcionalidades de interface: o Observatório Português propõe várias categorias de sites e gráficos para uma visão global. O Observatório descrito neste artigo possui objetivos e funcionalidades similares. Por hora, são verificados apenas sites governamentais com um modelo nacional de acessibilidade de forma automatizada, não oferecendo testes manuais. A interface não exibe gráficos e há apenas a categorização por Unidade Federativa. Apesar disso, há estatísticas globais e o nível de detalhes de cada página é maior. Por fim, a quantidade de páginas armazenadas é cerca de 230 vezes maior, aumentando a capacidade de contribuição da nossa ferramenta.

2.2 Modelos WCAG e eMAG e Software ASES

O modelo internacional WCAG, desenvolvido pelo consórcio W3C, define recomendações ou diretrizes de acessibilidade, que explicam como tornar o conteúdo Web acessível para todas as pessoas, auxiliando os administradores dos sítios.

O eMAG é basicamente a adaptação do WCAG para o contexto da web do governo brasileiro. Uma diferença conceitual importante é que as recomendações não estão divididas por níveis de prioridade, como é no WCAG, mas sim por área. São elas as seções de Marcação, Comportamento, Conteúdo/Informação, Apresentação/Design, Multimídia e Formulário.

Por tratarem-se de diretrizes para páginas governamentais, é ideal que todas aplicáveis ao contexto sejam seguidas. Por exemplo, se a página é uma área de contato, uma atenção especial deve ser dada às recomendações de formulário; se apresentar vídeo, serão destacadas as recomendações de multimídia.

Assim como no WCAG, o conceito de "recomendação" também está presente no eMAG, e compreende boas práticas de alto nível, que estão dentro de cada "seção". Os "critérios" [6], por sua vez, são regras implementáveis, ligadas a cada recomendação e que podem ser do tipo "erro" ou "aviso" exclusivamente.

Por fim, de maneira geral, a nota de avaliação é calculada pelo módulo eSelo [4], parte integrante do software avaliador ASES, através do somatório das notas atribuídas aos elementos do código fonte da página que estão em acordo com as recomendações e critérios

definidos pelo eMAG. Considerando que a nota é um número decimal de zero a cem, são definidas faixas de conformidade: [0, 70), [70, 85), [85, 95), [95, 100].

3 OBSERVATÓRIO DA ACESSIBILIDADE

3.1 Arquitetura

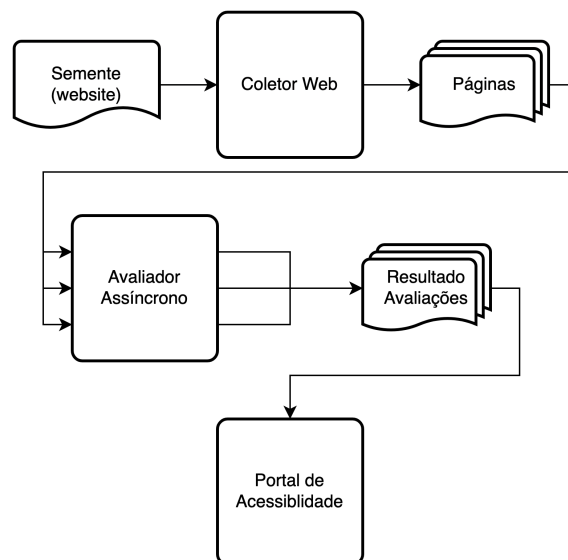


Figure 1: Arquitetura

3.1.1 Coletor Web. O Coletor Web é um módulo da ferramenta que extrai todos os endereços das páginas de um determinado site. É baseado nos conceitos de *Crawler* e *Scraper*, estratégias que percorrem recursivamente todos os links da página em profundidade.

3.1.2 Avaliador Assíncrono. O Avaliador Assíncrono é um módulo que gerencia o processamento de páginas coletadas pela ferramenta de avaliação de acessibilidade. O gerenciador instancia módulos ASES e define, levando em conta o número de páginas a serem analisadas, filas de processamento. Os arquivos JSON, resultantes do avaliador, são inseridos no banco de dados do portal.

3.1.3 Portal de Acessibilidade. O Portal de Acessibilidade é uma ferramenta Web que armazena e exibe as páginas analisadas pelo ASES e constrói indicadores globais.

3.2 Metodologia de Trabalho

Os sítios a serem avaliados partem de uma lista contendo candidatos pertencentes ao domínio alvo. A primeira etapa do processo consiste na coleta de todas as páginas contidas em cada um dos sítios candidatos por meio do componente Coletor Web. Nosso objetivo é ampliar ao máximo a cobertura do site, desconsiderando páginas muito profundas e pouco acessadas e otimizando os recursos computacionais. Ao final dessa etapa, também são realizados testes de disponibilidade, por meio da avaliação do status de resposta para as requisições HTTP, e filtro de compatibilidade de tipo de arquivo, eliminando arquivos que não estejam no formato HTML. A segunda

etapa consiste na avaliação das páginas coletadas pela ferramenta de análise de acessibilidade ASES, que sumariza os resultados da avaliação em uma “nota de acessibilidade”, além de apresentar o número de erros encontrados e sua localização no código fonte. Na terceira e última etapa, esses resultados são novamente filtrados e armazenados no banco de dados do Portal observador para serem agregados e visualizados.

3.3 Detalhes de Implementação

A implementação tem alguns requisitos a serem satisfeitos:

- páginas coletadas têm que estar disponíveis e funcionais, limitando a páginas HTML com código de retorno 200.
- as diretrizes constantes de arquivos como robots.txt têm que ser seguidas, eventualmente sendo necessário contatar o administrador do sítio.
- o conteúdo da página deve estar inteiramente em HTML, pois os scripts Javascript atualmente não são executados pelo Coletor.

Na versão atual, o processo de coleta é estático, ou seja, não recorrente, e para isso estamos utilizando a biblioteca Scrapy ¹. Futuramente, pretendemos realizar a migração para o arcabouço DynWebStats [8], que identifica páginas a serem coletadas e utiliza um escalonador sofisticado para realizar o acesso periódico às páginas armazenadas.

No *backend* do Portal de Acessibilidade, utilizamos o framework web Ruby On Rails ² e o SGBD não-relacional MongoDB ³. Os resultados da avaliação de acessibilidade foram mapeados em objetos utilizando ORM (*Object Relational Mapper*), facilitando a modelagem da interface de inserção dos dados. Para o *frontend*, foi utilizado o *framework* Bootstrap ⁴ e Javascript. Tanto o Avaliador Assíncrono quanto o Portal de Acessibilidade são executados em contêineres Docker ⁵ para maior facilidade e flexibilidade.

4 ESTUDO DE CASO: A WEB GOVERNAMENTAL BRASILEIRA

Nesta seção, apresentamos um estudo de caso de avaliação da ferramenta na Web governamental brasileira, utilizando uma semente de 8 mil sítios, avaliados na profundidade máxima igual a 3.

Ao acessar o Portal, o usuário tem acesso a um breve resumo dos objetivos da ferramenta e visualiza o número de sites e de páginas armazenadas. Em seguida, é apresentada uma tabela que agrega as páginas de acordo com as faixas de conformidade apresentadas por ASES, como ilustrado na Figura 2.

Nas abas de Sites Verificados e 10 Melhores Sites estão as listagens de sites, com buscas por endereço e título e filtros de notas e número de páginas. Ao clicar em um site, por exemplo www.into.saude.gov.br, observa-se em destaque um sumário de informações globais do site e as informações das análises de cada uma das páginas pertencentes. Ao clicar em uma página, é possível ver as recomendações por seção, bem como os critérios e o código fonte, como na Figura 3.

¹<https://scrapy.org/>
²<https://rubyonrails.org/>
³<https://www.mongodb.com/>
⁴<https://getbootstrap.com/>
⁵<https://docker.com>

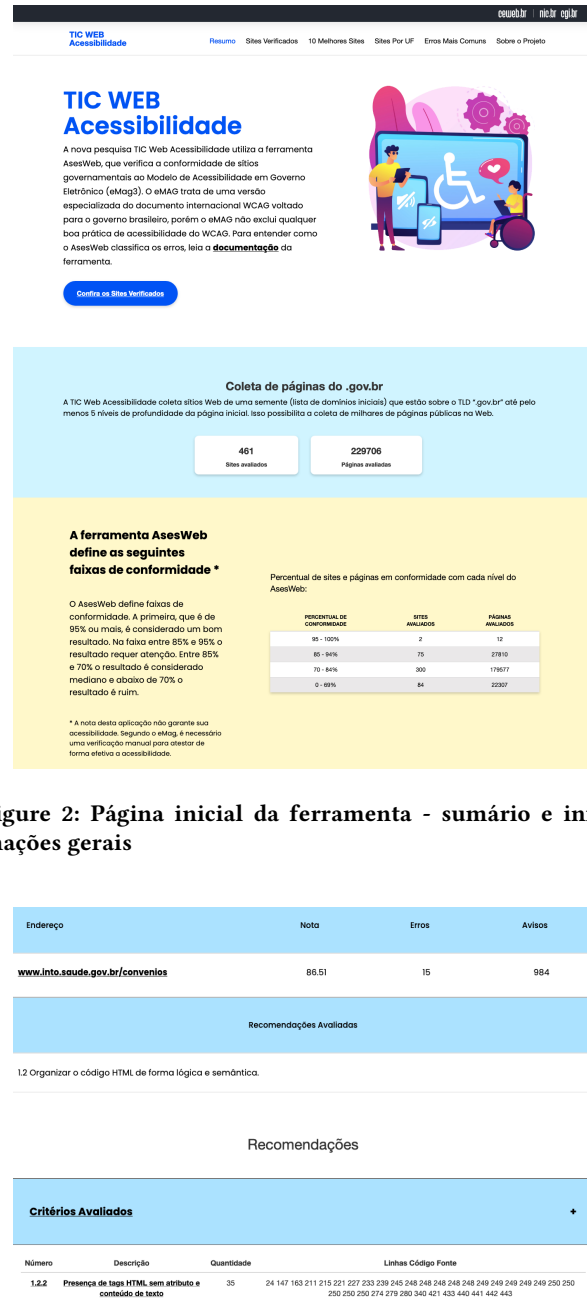


Figure 2: Página inicial da ferramenta - sumário e informações gerais



Figure 3: Exemplo da visualização dos critérios de uma recomendação, como sugestão de correção da página.

A Aba Sites por UF agrega os sites de acordo com as Unidades Federativas, permitindo ao usuário analisar as notas de acessibilidade em um contexto geográfico.

Por fim, a aba de Erros Mais Comuns, como mostra a Figura 4, apresenta a ocorrência das recomendações do eMAG por seção, além da descrição e do somatório de cada erro encontrado por sites e por páginas. No caso, a seção de Marcação está em destaque,

uma vez que esses erros são cometidos por quase todas as páginas verificadas.

Número	Descrição	Total de Erros	Total de Sites	Total de Páginas
1.1	Respeitar os Padrões Web	7.241.217	415	227.140
1.2	Organizar o código HTML de forma lógica e semântica	1.381.146	434	228.903
1.3	Utilizar corretamente os níveis de cabeçalho	430.272	450	229.368
1.5	Fornecer âncoras para ir direto a um bloco de conteúdo	440.203	437	225.795
1.6	Não utilizar tabelas para diagramação	6.721	76	37.756
1.7	Separar links adjacentes	325.269	195	164.160

Figure 4: Listagem dos Erros Mais Comuns, com destaque para a seção de erros do tipo Marcação

5 CONSIDERAÇÕES FINAIS

Este artigo apresentou uma ferramenta para avaliação da acessibilidade da Web, contendo várias funcionalidades. Essa ferramenta, aqui denominada “Observatório da acessibilidade da Web brasileira” está publicada sob a Licença Creative Commons 4.0 (CC BY-SA 4.0) e está disponível para acesso no endereço Web <http://acessibilidade.speed.dcc.ufmg.br/ases/resumo>, com a seguinte autenticação de demonstração: usuário “demo” e senha “demo”.

Acreditamos que os resultados obtidos foram positivos e, considerando o caso de uso da Web brasileira, serão úteis para cidadãos interessados em verificar o nível de acessibilidade nos sítios públicos e para desenvolvedores e administradores tomarem conhecimento das boas práticas e de como solucionar possíveis erros.

Futuramente pretende-se adicionar à semente, que atualmente compreende apenas o gov.br, sites dos domínios jus.br, mp.br e mil.br. Pretende-se também evoluir a coleta estática das páginas para um mecanismo dinâmico e inteligente já desenvolvido, o arcabouço de coleta dinâmica DynWebStats. Além do ASES, novos analisadores com modelos diferentes podem ser adicionados, como o WCAG. Finalmente, no escopo das funcionalidades da interface de usuário, ampliaremos a categorização dos sites, incluindo instituições educacionais, municípios, ministérios, entre outros. Também possibilitaremos a sugestão de sites e páginas pelos usuários.

REFERENCES

- [1] Section 508. 2018. Applicability and Conformance Requirements. Retrieved 30 de Setembro de 2021 from <https://www.section508.gov/develop/applicability-conformance/>
- [2] Portuguese Observatory of Web Accessibility Administrative Modernization Agency. 2020. Observatório Português da Acessibilidade Web. Retrieved 30 de Setembro de 2021 from <https://observatorio.acessibilidade.gov.pt/>
- [3] Tim Berners-Lee. 1991. World Wide Web. Retrieved 30 de Setembro de 2021 from <http://info.cern.ch/hypertext/WWW/TheProject.html>
- [4] Programa de Governo Eletrônico Brasileiro. 2015. Métricas para avaliação de acessibilidade virtual. Retrieved 30 de Setembro de 2021 from https://gitlab.sgd.nuvem.gov.br/asesweb/e-selo/-/blob/master/Documentacao/M%C3%A9tricas_ASES_eSelo_Checklist_Junho2016.pdf
- [5] Programa de Governo Eletrônico Brasileiro. 2017. Avaliador e Simulador de Acessibilidade em Sítios. Retrieved 30 de Setembro de 2021 from <https://asesweb.governoeletronico.gov.br/>
- [6] Programa de Governo Eletrônico Brasileiro. 2017. Critérios de Sucesso. Retrieved 30 de Setembro de 2021 from <https://asesweb.governoeletronico.gov.br/criteriosSucesso>

- [7] Programa de Governo Eletrônico Brasileiro. 2021. eMAG - Modelo de Acessibilidade em Governo Eletrônico. Retrieved 30 de Setembro de 2021 from <http://emag.governoeletronico.gov.br/>
- [8] Israel Guerra, Wagner Meira, Adriano César Machado Pereira, Diogo Marques Santa, Vagner Diniz, Heitor Ganzeli, Marcelo Pitta, and Alexandre Barbosa. 2016. DynWebStats: A Framework for Determining Dynamic and Up-to-Date Web Indicators. In *Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web (Teresina, Piauí State, Brazil) (Webmedia '16)*. Association for Computing Machinery, New York, NY, USA, 247–254. <https://doi.org/10.1145/2976796.2976857>
- [9] Dilma Rousseff. 2015. Lei Brasileira de Inclusão da Pessoa com Deficiência. Retrieved 30 de Setembro de 2021 from http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm
- [10] W3C. 1989. História da WWW. Retrieved 30 de Setembro de 2021 from <https://w3.org/History/1989/proposal.html>
- [11] W3C. 1997. World Wide Web Consortium (W3C) Launches International Web Accessibility Initiative. Retrieved 30 de Setembro de 2021 from <https://www.w3.org/Press/WAI-Launch.html>
- [12] W3C. 2021. Web Content Accessibility Guidelines (WCAG). Retrieved 30 de Setembro de 2021 from <https://www.w3.org/WAI/standards-guidelines/wcag/>