

Fecha de recepción: marzo 2023
Fecha de aceptación: abril 2023
Versión final: mayo 2023

Evaluación de la accesibilidad web: oportunidades con inteligencia artificial y aprendizaje automático

Edson Rufino de Souza^(*)

Resumo: A acessibilidade na web é essencial para a plena participação de pessoas com ou sem deficiência na sociedade. No entanto, a maioria dos sites não atende aos requisitos mínimos de acessibilidade. A avaliação da acessibilidade é uma das etapas para tornar os sites mais acessíveis. Apresentamos desafios relacionados e discutimos possibilidades atuais do uso de inteligência artificial e aprendizado de máquina para melhorar a acessibilidade de sites na Internet.

Palavras-chave: acessibilidade - inteligência artificial - aprendizado de máquina

[Resumos em inglês e espanhol na página 254]

^(*) Diseñador y Doctor en Design. Profesor adjunto del Curso de Diseño de la Universidad Federal do Espírito Santo – UFES (Vitória, ES, Brasil). Coordinador del Laboratorio del Diseño y Interacción (LDI/SEAD/UFES) y del Grupo de Investigación en Formalizaciones Matemáticas de la Cognición y Diseño (Forma/UFES). Miembro del grupo de trabajo en Accessibilidad Web del W3C Brasil.

Introdução

É inegável que o meio digital adquiriu papel preponderante na vida cotidiana atual. Aplicativos, sítios eletrônicos, jogos e outros artefatos em mídia eletrônica são elementos fundamentais no cotidiano de pessoas das mais variadas faixas etárias e classes sociais. A internet especificamente converge trabalho, negócios, serviços, pesquisa, entretenimento e parte significativa das nossas relações sociais.

A explosão do uso de dispositivos móveis e do avanço do uso da grande rede especialmente no meio doméstico tem acentuado ainda mais a presença constante e diária da web na vida das pessoas. Dados recentes demonstram que no Brasil há mais smartphones ativos do que pessoas (Meirelles, 2022).

Segundo o ICSID (2000), “design é o fator central da humanização inovadora das tecnologias e o fator crucial das trocas econômicas e culturais”. Nesta humanização, entendemos que a inclusão deveria estar incluída como diretriz para os projetos que desenvolvemos. Em 1997, Tim Berners-Lee, o criador da web, afirmou: “O poder da web está na sua universalidade. O acesso por todos independentemente de deficiência é um aspecto essencial”. Pode-se dizer que não há o tal alcance mundial sem que haja acesso por diferentes pessoas, independentemente de suas características e necessidades específicas.

A grande maioria dos sítios eletrônicos disponíveis pode atender uma gama diversificada de usuários, desde que sejam projetados e desenvolvidos para que sejam mais acessíveis. Isto significaria ter um número maior de pessoas se beneficiando dos recursos existentes hoje na internet. O que inclui pessoas com os perfis diversos, inclusive utilizando diferentes dispositivos e em contextos variados de uso.

Neste sentido, uma conquista fundamental foi a adoção da Convenção Internacional sobre os Direitos das Pessoas com Deficiência, e sua posterior ratificação pelos diferentes países ao redor do mundo. No momento, a Convenção e seu Protocolo Facultativo tem assinatura de 164 países. Foi promulgada no Brasil pelo Decreto Federal nº 6.949, de 25 de agosto de 2009.

O texto traz a perspectiva fundamental de deixar de ver as pessoas com deficiência como “objetos” de caridade, tratamento médico e proteção social para vê-las como “sujeitos” com direitos, capazes de reivindicar esses direitos e tomar decisões por suas vidas com base em seu consentimento livre e informado, bem como sendo membros ativos da sociedade (Nações Unidas, 2022).

A Convenção determina no seu Artigo 9, “Acessibilidade”:

2. Os Estados Partes deverão também tomar medidas apropriadas para: (...)
 - g. Promover o acesso de pessoas com deficiência a novos sistemas e tecnologias da informação e comunicação, inclusive à internet; e
 - h. Promover o desenho, o desenvolvimento, a produção e a disseminação de sistemas e tecnologias de informação e comunicação em fase inicial, a fim de que estes sistemas e tecnologias se tornem acessíveis a um custo mínimo.

No âmbito local, diversas legislações nacionais têm surgido dentro do escopo da acessibilidade de sítios eletrônicos. No Brasil, a Lei nº 13.146, de 6 de julho de 2015, que institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência) definiu que:

“Art. 63: É obrigatória a acessibilidade nos sítios da internet mantidos por empresas com sede ou representação comercial no País ou por órgãos de governo, para uso da pessoa com deficiência, garantindo-lhe acesso às informações disponíveis, conforme as melhores práticas e diretrizes de acessibilidade adotadas internacionalmente.” (Brasil, 2015)

Em todo o mundo, estima-se que há mais de um bilhão de pessoas com algum tipo de deficiência, segundo a Organização Mundial da Saúde (2021). Os dados brasileiros mais recentes, da Pesquisa Nacional de Saúde de 2019 (PNS), também realizada pelo IBGE, apresentam o número de mais de 17 milhões de pessoas com deficiência no Brasil (IBGE, 2021). Com quantitativo tão significativo de pessoas com deficiência no Brasil e no mundo, fica ainda mais evidente a necessidade e a urgência por mais acessibilidade em todos os meios. No meio digital, a acessibilidade é constantemente relacionada ao design e desenvolvimento de sistemas interativos mais inclusivos. Contudo, existem definições e visões diferentes entre si do que significa a acessibilidade no meio digital. Pode-se então deduzir que, consequentemente, existem maneiras diversas para avaliar e projetar para a acessibilidade, que vem sendo discutidas pelos pesquisadores da área nos últimos anos (Souza, 2016). Por exemplo, a ISO 9241-171 (2008) define acessibilidade como a “usabilidade de um produto, serviço, ambiente ou facilidade para pessoas com a mais ampla variedade de capacidades”. Já para o World Wide Web Consortium (W3C), “acessibilidade web significa que pessoas com deficiência possam usar a web. Em outras palavras, elas devem poder perceber, compreender, navegar e interagir com a web, e assim devem poder contribuir com a web” (W3C, 2022).

Pode-se perceber focos bastante distintos nestas definições. A definição do W3C prioriza de forma direta as pessoas com deficiência, enquanto a definição da norma ISO apresenta um direcionamento menos específico para diferentes pessoas.

Existem diferentes propostas para a avaliação de acessibilidade de sítios eletrônicos no contexto da Web. Em relação às técnicas existentes para se avaliar acessibilidade no meio digital, Yesilada et al (2009) afirmam que estas se dividem basicamente em cinco categorias: técnicas de inspeção (avaliação por especialistas); testes automáticos (com uso de software); técnicas de screening; avaliações subjetivas; e testes com usuários.

O problema do especialista avaliador de acessibilidade

Embora muitos problemas só sejam descobertos com a avaliação com usuários (DRC, 2004; Souza, 2005; Power et al, 2012), existe em geral dificuldade em recrutar pessoas com diferentes tipos de deficiência e em contextos reais de utilização. Por isso, as avaliações automáticas e as técnicas de inspeção são muito utilizadas para avaliação de acessibilidade. Waller e Clarkson (2009) apresentam as técnicas de avaliação por inspeção como “Expert Appraisal”, que seria o “exame feito por perito especialista de um produto ou serviço por alguém que tem treinamento profissional ou experiência para fazer um julgamento informado sobre o projeto”. Os autores complementam ainda que esta técnica “pode identificar possíveis causas de exclusão do projeto, sugerir melhorias para reduzir esta exclusão e aumentar a satisfação do usuário”.

Brajnik et al (2011) demonstraram que avaliadores experientes têm maior efetividade e confiabilidade ao buscar problemas em um dado conjunto de páginas Web do que avalia-

dores novatos treinados para este fim. Em seu estudo, os pesquisadores observaram que três especialistas mais experientes foram capazes de encontrar um número de problemas que só 14 avaliadores menos experientes conseguiram encontrar.

Waller e Clarkson (2009) destacam ainda que, idealmente, o “especialista” não deveria ser enviesado por um envolvimento anterior com o projeto porque a familiaridade com qualquer produto ou tarefa faz com que esta pareça mais simples e fácil, o que pode dificultar que certos problemas fossem encontrados.

O efeito do avaliador é definido por Hornbeak e Frokjear (2008) como a observação de avaliadores de usabilidade em condições semelhantes poderem identificar conjuntos substancialmente diferentes de problemas de usabilidade.

Esta conclusão foi gerada pelos pesquisadores a partir de um levantamento de estudos anteriores (Jacobsen et al, 1998; Hertzum & Jacobsen, 1999; Hertzum & Jacobsen, 2001; Hertzum et al, 2002) que verificaram a ocorrência de diferenças significativas entre os resultados apresentados pelos diferentes avaliadores que participaram dos estudos.

Nestes estudos, foram empregadas técnicas diversas de avaliação de usabilidade, com ou sem a presença do usuário, como simulação cognitiva (ou percurso cognitivo), avaliação heurística e testes com usuários usando o protocolo de “pensar em voz alta”, mas onde estava presente a figura do avaliador. De acordo com Hertzum & Jacobsen (2001), a concordância média entre avaliadores de usabilidade do mesmo sistema, utilizando a mesma técnica varia entre 5% a 65%.

Alguns pesquisadores investigaram os fatores envolvidos no efeito do avaliador. Hertzum et al (2002) afirmam que um procedimento de avaliação vago pode fazer diferentes avaliadores se concentrarem em coisas diferentes durante a avaliação.

No mesmo estudo, Hertzum et al (2002) concluem que a principal causa para o efeito do avaliador é que a avaliação de usabilidade é uma atividade cognitiva que exige que os avaliadores exerçam julgamento. Hornbeak & Frokjear (2008) afirmam ainda que os avaliadores podem ocasionalmente deixar de observar a evidência de um determinado problema.

Mesmo com protocolos de avaliação e com indivíduos avaliadores com conhecimento e experiência em diferentes aspectos relacionados à acessibilidade, a avaliação de acessibilidade por especialistas em acessibilidade encontra dificuldades para ter resultados consistentes mesmo entre pessoas com alto nível de expertise em diferentes aspectos relacionados à acessibilidade (Souza, 2015).

Na avaliação com usuários, a principal diferença é que o avaliador especialista não verifica sozinho um sítio de acordo com uma lista de critérios de sucesso: o avaliador tem de identificar as questões de acessibilidade de um sítio eletrônico a partir dos dados obtidos nas sessões dos testes, que podem ser problemas diretamente observados na interação ou a partir de comportamentos observados e das reações dos participantes durante o estudo. Por isso, fontes de viés podem ocorrer antes, durante e após o teste com usuários (Aizpuru et al, 2014).

Entre esses vieses, também se deve considerar o chamado efeito do avaliador (Jacobsen et al, 1998). Neste estudo, quatro avaliadores analisaram gravações em vídeo de quatro sessões de testes de usabilidade com usuários. Somente 20% dos problemas únicos foram

detectados por todos os quatro avaliadores, enquanto 46% foram encontrados por um único avaliador.

Dentre os problemas considerados mais graves, a distância entre os avaliadores foi menor, com 41% dos problemas sendo encontrados por todos os quatro avaliadores e 22% por um único avaliador. Em ambos os casos, percebe-se impacto significativo do avaliador na interpretação dos dados dos testes.

Acessibilidade, conformidade e avaliação automática

No contexto da acessibilidade, as técnicas de inspeção são baseadas na verificação por um avaliador de se uma página web é ou não acessível. A técnica de inspeção mais amplamente utilizada para avaliação de acessibilidade é a avaliação de conformidade, onde um avaliador usa como base um conjunto de diretrizes estabelecidas a partir de possíveis problemas de acessibilidade, e decide se uma página ou sítio eletrônico está ou não em conformidade com estes requisitos (Yesilada et al, 2009).

Por conta da acessibilidade se constituir num direito, existe a necessidade de se definir parâmetros comuns e inequívocos de como esta pode ser avaliada. As Diretrizes de Acessibilidade de Conteúdo para a Web, ou Web Content Accessibility Guidelines (WCAG) são um documento de caráter normativo que se descreve como meio de definir a forma como tornar o conteúdo da Web mais acessível a pessoas com deficiências (World Wide Web Consortium [W3C], 2018).

O WCAG 2.1 (W3C, 2018), versão atual do documento, é uma atualização do WCAG 2.0 (W3C, 2008). A estrutura básica do WCAG é definida por quatro princípios, que compõem as categorias em que se dividem doze diretrizes de acessibilidade. Estas se subdividem em critérios de sucesso testáveis (W3C, 2018). Estes últimos serão os pontos a serem verificados em uma avaliação.

Contudo, o documento tem limitações. É estabelecido no documento que o WCAG ainda não consegue atender às necessidades de pessoas com todos os tipos, graus e combinações de deficiência. Houve desafios significativos na definição de critérios adicionais para abordar deficiências intelectuais, bem como transtornos de linguagem e de aprendizado, considerando o cronograma de desenvolvimento das diretrizes, a dificuldade em estabelecer consenso sobre testabilidade, implementabilidade e considerações internacionais e culturais nas diferentes propostas (W3C, 2018).

Como definido no próprio documento, o WCAG foi projetado para ser aplicado de forma ampla a diferentes tecnologias da Web, agora e no futuro, e para ser testado com uma combinação de testes automatizados e humanos (W3C, 2018). Esta premissa de ter que juntar a abordagem humana da que é feita por ferramentas de software se baseia nas dificuldades de as ferramentas de teste verificarem de forma efetiva diversos critérios de sucesso do WCAG que atualmente só podem ser avaliados por técnicas de inspeção, por pessoas com expertise em acessibilidade (Yesilada et al, 2009).

Por exemplo, o critério de sucesso 3.1.1 do WCAG 2.1, “Idioma da Página”, define que “O idioma humano predefinido de cada página da Web pode ser determinado de forma programática”, o que pode ser verificado de forma bastante simples por ferramentas de teste de acessibilidade atualmente disponíveis, sem necessidade de checagem humana. Por outro lado, o critério de sucesso (1.1.1) “Conteúdo Não-Textual” define que “todo o conteúdo não textual apresentado ao usuário tem uma alternativa em texto que serve a uma finalidade equivalente”. A verificação de se um dado texto alternativo proposto pode ser considerado como um equivalente textual para uma imagem no contexto onde foi inserida normalmente só é realizada por seres humanos com conhecimento em acessibilidade, não sendo avaliada por ferramentas de teste tradicionais.

Números da acessibilidade web

Muito infelizmente, mesmo com a difusão das diretrizes de acessibilidade desde 1999, a maioria dos sítios eletrônicos não atende às boas práticas propostas em documentos existentes de diretrizes de acessibilidade. Ao longo dos tempos, diversos estudos verificaram que a quantidade de páginas que atendem aos requisitos de acessibilidade web estão muito abaixo do que seria ideal.

Em 2010, em avaliação automatizada de uma coleção de quase 30 milhões de páginas da Web fornecida pelo Web Archive português, foi encontrada alta variabilidade em relação ao nível de acessibilidade das páginas da Web e poucas páginas atingiram altos níveis de acessibilidade. Os resultados obtidos mostraram também que existe uma correlação entre acessibilidade e complexidade (em número de elementos HTML) de uma página da Web. (Lopes et al, 2010).

Outros estudos recentes confirmam a tendência de números baixos de páginas acessíveis. O estudo “WebAIM Million” consistiu de uma avaliação de acessibilidade das páginas iniciais dos 1.000.000 principais sites (definidos pelo cruzamento de estatísticas de três serviços distintos). A avaliação foi realizada usando a API autônoma WAVE. 96,8% das páginas iniciais dos sites definidos tinham falhas em relação ao WCAG 2 (WebAIM, 2022).

No Brasil, estudo confirmou esta tendência em relação aos sites brasileiros. 14 milhões de sites brasileiros ativos foram avaliados, e menos de 1% passou nos testes de acessibilidade. No caso dos sites governamentais, esse percentual cai para 0,34%. A maioria dos endereços on-line do País, ou 93,79%, encontra-se em uma zona cinzenta: apresentou falha em algum dos testes realizados, mas pontuou positivamente em outros (BigDataCorp & MWPT, 2022).

Pelo volume de avaliações, os três estudos citados acima só foram possíveis por serem avaliações a partir de ferramentas automáticas (software), que avaliam apenas parte dos critérios de sucesso do WCAG. Assim, os resultados muito provavelmente seriam piores se os critérios de sucesso checados apenas por humanos fossem incluídos. Ademais, os dois últimos estudos são anuais e estão ambos em sua quarta edição, e têm apresentado resultados consistentes quanto à grande maioria das páginas estarem inadequadas em relação à acessibilidade.

Destaca-se ainda que, com o número crescente de sites sendo criados a cada dia somado ao aumento da complexidade dos sítios eletrônicos (WebAIM, 2022), ter uma internet que atenda aos requisitos de acessibilidade é um desafio ainda sem solução.

Inteligência artificial e aprendizado de máquina

O termo “Inteligência Artificial” (IA) foi idealizado por McCarthy em 1956, e usado pela primeira vez no convite para uma conferência de verão em que se estudaria de forma dedicada a automação de tarefas e a “inteligência” das máquinas aplicada a tarefas até então só pensadas a ser realizadas por seres humanos.

Contudo, a questão de como máquinas poderiam emular de alguma forma o pensamento humano é algo vislumbrado e estudado há mais tempo. Seis anos antes, Alan Turing se tornou famoso por seu artigo “Computer Machinery and Intelligence” (1950). Neste estudo, a partir da pergunta sobre se as máquinas podem pensar, ele trabalhou sob a suposição de que máquinas podem adquirir o domínio da linguagem de forma a serem confundidos com seres humanos.

A partir desta premissa, Turing questionou se máquinas teriam sucesso no “jogo da imitação”. Resumidamente, um jogador humano participa de uma conversa, em linguagem natural, com outro humano e uma máquina que tenha sido projetada para gerar respostas capazes de ser confundidas com as de uma pessoa. Estando os participantes todos separados, um juiz teria que julgar se a máquina passou no teste de se distinguir dos humanos participantes do diálogo. O teste avalia as respostas não em relação à exatidão das mesmas, mas sim o quanto estas se aproximam das respostas que poderiam ser dadas por uma pessoa real.

Voltando à pesquisa de Dartmouth, a proposta de estudo partia da premissa de que “cada aspecto do aprendizado ou qualquer outra característica da inteligência pode, em princípio, ser descrito com tanta precisão que uma máquina pode ser feita para simular os mesmos”. Assim, o objetivo era uma “tentativa de descobrir como fazer as máquinas usarem a linguagem, formar abstrações e conceitos, resolver tipos de problemas agora reservados aos humanos e se aprimorarem”.

Esta premissa se aproxima mais da visão contemporânea da inteligência artificial, em que microprocessos como reconhecimento de imagens ou de voz são analisados e transformados em algoritmos, a fim de que possam ser processados por sistemas.

Não existe consenso sobre uma definição do que é inteligência artificial. Russel e Norvig (2016), autores de um livro basilar sobre o tema, definem a inteligência artificial como a concepção e a construção de agentes inteligentes que podem receber percepções do ambiente e executar ações que afetam esse mesmo ambiente.

Ainda segundo os autores, diferentemente de áreas como a Filosofia ou a Psicologia, o foco da IA não está apenas em entender o pensamento humano, mas também em construir entidades capazes de fazê-lo. Assim, a IA busca a criação de programas que possam emular nosso raciocínio, tomar decisões e definir ações (Russel & Norvig, 2016).

O aprendizado de máquina ou aprendizagem automática, é atualmente um dos principais ramos do campo de inteligência artificial para o desenvolvimento de abordagens para realização de tarefas antes só realizadas com o apoio direto da cognição humana, junto a outros campos. Aprendizado de máquina, aprendizado profundo e redes neurais são todos subcampos da inteligência artificial. No entanto, as redes neurais são, na verdade, um subcampo do aprendizado de máquina, e o aprendizado profundo é um subcampo das redes neurais (Kavlakoglu, 2022).

Problema e objetivo da pesquisa

A partir do que contexto acima estabelecido, esta pesquisa concentra-se em maneiras de tornar mais efetiva a avaliação por especialistas com o uso de ferramentas, a partir da sistematização do método de inspeção e da aplicação de agentes inteligentes que deem apoio à interpretação e análise por seres humanos.

Desta forma, esta pesquisa tem caráter exploratório e toma como base o questionamento sobre a viabilidade e a efetividade da interpretação por agentes baseados em inteligência artificial e/ou aprendizado de máquina de aspectos da avaliação de acessibilidade web, normalmente avaliados apenas por humanos especialistas em acessibilidade.

Assim, o objetivo da pesquisa é identificar possibilidades de melhorias na avaliação de acessibilidade web com base no WCAG utilizando de inteligência artificial (IA) ou aprendizado de máquina (AM).

Metodología

Com este fim, este estudo teve como foco uma revisão de literatura a partir da técnica de mapeamento sistemático de literatura (MSL) (PETERSEN et al 2010; 2015), com o foco em propostas de aplicação de inteligência artificial na avaliação de acessibilidade em contexto digital.

Os estudos de MSL são usados para estruturar uma área de pesquisa e, por isso, não tem foco apenas em estudos que apresentem evidências. Por outro lado, as revisões sistemáticas de literatura (RSL) se concentram na coleta de artigos num escopo bastante específico e na síntese de evidências (Petersen et al, 2015).

Segundo Kitchenham (2010), questões de pesquisa em MSL são gerais, pois visam descobrir tendências de pesquisa (por exemplo, tendências de publicação ao longo do tempo, tópicos abordados na literatura de um dado campo). Por outro lado, as RSL visam agragar evidências e, portanto, um objetivo muito específico deve ser formulado (por exemplo, se um tipo específico de intervenção é praticamente útil para a indústria).

Inicialmente, os MSL eram recomendados principalmente para áreas de pesquisa em engenharia de software em que faltavam estudos primários relevantes e de alta qualidade

(Kitchenham et al, 2007). E por isso são adequados para pesquisas em temas mais recentes como o desta pesquisa.

Atualmente os MSL são recomendados para fornecer uma visão geral de uma área de pesquisa por meio da classificação e contagem de contribuições em relação às categorias dessa classificação. Normalmente envolve enumerar quais tópicos do tema pesquisado foram abordados na literatura e onde os trabalhos relacionados têm sido publicados (Petersen et al, 2015).

As fases do processo do MSL segundo Petersen et al (2008):

Passos do processo de MSL - Petersen et al (2008)	Resultados
definição de questão de pesquisa	escopo de revisão
realizar busca	todos os artigos
trigam de artigos	artigos relevantes
definição de palavras-chave usando resumos	esquema de classificação
extração de dados e processo de mapeamento	mapa sistemático

Tabela 1: passos do processo de MSL - Petersen et al (2008)

Desenvolvimento da pesquisa

A partir do problema e do objetivo, foram estabelecidas as seguintes questões de pesquisa (QP):

- QP1: quais recursos e/ou técnicas de IA/AM vem sendo testados na avaliação de acessibilidade web pela revisão de conformidade com o WCAG?
- QP2: Que tipos de problema de avaliação de acessibilidade web se buscou resolver com a aplicação de IA e/ou AM?
- QP3: Os estudos apresentam evidências sobre vantagens da aplicação de IA / AM em relação a abordagens tradicionais de avaliação de acessibilidade por software?
- QP4: Onde e quando os estudos encontrados foram publicados?

A partir da definição das questões de pesquisa, passou-se à estratégia de busca. Pelas características do MSL, e pela necessidade de se acessar todos os tipos de estudos disponíveis, optou-se por realizar a pesquisa diretamente pelo Google Scholar, em vez de escolher bases acadêmicas específicas. Isto foi feito por três motivos:

- a. para evitar a possibilidade de viés editorial, o que pode estar presente utilizando-se a busca específica de um dado conjunto de bases de dados, e incluir a chamada “literatura cinza”; e

- b. pela busca de maior alcance que o mecanismo de busca apresenta; e
- c. pelo fato de o Google Scholar (GS) ser atualmente a mais abrangente ferramenta de busca acadêmica.

Segundo Gusenbauer (2018), o GS, com 389 milhões de registros, era o mecanismo de busca acadêmico mais abrangente, considerando o ano de 2018. Além disso, estudo de Martín-Martín et al (2018) demonstrou que o GS encontrou consistentemente a maior porcentagem de citações em todas as áreas de conhecimento (93% –96%), muito à frente comparativamente com mecanismos de bases como Scopus (35–77%) e Web of Science (27–73%). GS encontrou quase todas as citações de WoS (95%) e Scopus (92%). A maioria das citações encontradas apenas por GS foram de fontes não periódicas (48– 65%), incluindo teses, livros, artigos de conferências e materiais não publicados. Os resultados sugerem que em todas as áreas os dados de citação GS são essencialmente um superconjunto de WoS e Scopus, com cobertura extra substancial.

Definido o mecanismo de busca, passou-se à definição das strings de pesquisa, definidas pelo conjunto de palavras-chave a serem usadas nas pesquisas. Optou-se por realizar a busca apenas em língua inglesa. As buscas foram feitas em fevereiro de 2022 e, considerando a novidade do tema, o escopo de tempo foi de trabalhos de 2018 a 2021. Foram definidas as seguintes strings de busca, diferenciadas pelo uso de “evaluation” ou “assessment”, ou de “artificial intelligence” ou “machine learning”.

- +”web accessibility” +”wcag” +”conformance” +”evaluation” +”artificial intelligence”
- +”web accessibility” +”wcag” +”conformance” +”evaluation” +”machine learning”
- +”web accessibility” +”wcag” +”conformance” +”assessment” +”artificial intelligence”
- +”web accessibility” +”wcag” +”conformance” +”evaluation” +”machine learning”

Os artigos encontrados a partir do conjunto de strings de busca só são considerados relevantes após passarem pela triagem de critérios de inclusão predefinidos. São eles:

- a. O resumo ou o título do trabalho do artigo cita resultados ou propostas de aplicação de técnicas/recursos de IA e AM?
- b. O resumo ou o título do trabalho citam resultados ou de propostas concretas de avaliação de acessibilidade web?
- c. O trabalho está disponível de forma completa gratuitamente sem necessidade de compra, pela rede convencional ou pela rede universitária?

Resultados

As buscas foram realizadas em fevereiro de 2022, a partir das strings de pesquisa predefinidas. Foram encontrados 117 trabalhos, dos quais 20 atenderam aos critérios de inclusão e exclusão a partir do título e do resumo dos trabalhos, e foram selecionados para leitura completa. A leitura completa busca confirmar o entendimento de que os artigos cumprem com os dois primeiros critérios de inclusão e, para os confirmados, analisa os trabalhos de acordo com base nas quatro questões de pesquisa predefinidas. A figura 1 apresenta a planilha utilizada para reanálise dos critérios de inclusão e exclusão na leitura completa.

Figura 1: Tabela de triagem dos artigos pré-selecionados após leitura completa.

Dos 20 trabalhos pré-selecionados, 7 trabalhos foram selecionados de acordo com os critérios estabelecidos, e são descritos na tabela a seguir, a partir das questões de pesquisa predefinidas. São apresentados os artigos selecionados que compõem o mapa relativo ao escopo desta pesquisa:

referência do trabalho	QP1: quais recursos e/ ou técnicas de IA/AM vem sendo testados na avaliação de acessibilidade web pela revisão de conformidade com o WCAG?	QP2: Que tipos de problema de avaliação de acessibilidade web se buscou resolver com a aplicação de IA e/ ou AM?	QP3: Os estudos apresentam evidências sobre vantagens da aplicação de IA / AM em relação a abordagens tradicionais de avaliação de acessibilidade por software?	QP4: Onde e quando os estudos encontrados foram publicados?
1 Dogra, H. (2020). A Framework for an automatic evaluation of image description based on an image accessibility guideline (Master's thesis, OsloMet-storbyuniversitetet).	Aprendizado de máquina usado para construir modelo que prevê o quanto acessível é uma descrição de imagem em relação às diretrizes do NCAM.	Verificar se um texto alternativo para uma imagem é adequado	A avaliação qualitativa da estrutura proposta pelos especialistas valida seu desempenho e eficácia, considerando-a comparável à avaliação manual de textos alternativos por especialistas em acessibilidade.	Tese de Mestrado em Universal Design of ICT - OsloMet, Noruega
2 Haider, W. (2021). Automatically Detecting Table Headings In Web Pages With Machine Learning (Master's thesis, Middle East Technical University).	Detecção das tabelas e dos seus cabeçalhos em páginas Web usando recursos das páginas renderizadas com algoritmos de aprendizado de máquina.	Identificar a estrutura da tabela e se esta está marcada corretamente no HTML. Ferramentas comuns apenas verificam o HTML para identificar se no elemento <table> a primeira linha está marcada como cabeçalho. As demais verificações deste tipo de marcação são feitas por seres humanos.	O trabalho conseguiu identificar tabelas em páginas e seus cabeçalhos em página da Web com precisão superior a trabalhos anteriores.	Tese de Mestrado (2021) em Computer Engineering - Middle East Technical University, Northern Cyprus Campus, Turquia
3 Rodriguez, G. (2020). Improving Web Accessibility Through Suggestions Using Serverless Architecture (Doctoral dissertation, California State University, Northridge).	Sistema de análise de texto e sistema de recomendações por chatbot	Informações para o desenvolvedor de conteúdo da Web com base em experiências anteriores de usuários finais para melhorar a acessibilidade em um determinado site.	A autora argumenta que houve ganhos no método de classificação de relatos de problemas de usuários finais de acordo com as diretrizes do WCAG. E encontrou desafios com a classificação de textos curtos, que se tornam mais conversacionais e menos precisos, que precisam de maior volume de dados de treinamento.	Dissertação de Doutorado (2020) em Software Engineering - California State University, Northridge, Estados Unidos

>> continúa

referência do trabalho	QP1: quais recursos e/ ou técnicas de IA/AM vêm sendo testados na avaliação de acessibilidade web pela revisão de conformidade com o WCAG?	QP2: Que tipos de problema de avaliação de acessibilidade web se buscou resolver com a aplicação de IA e/ ou AM?	QP3: Os estudos apresentam evidências sobre vantagens da aplicação de IA / AM em relação a abordagens tradicionais de avaliação de acessibilidade por software?	QP4: Onde e quando os estudos encontrados foram publicados?
4 Morato, J., Campillo, A., Sanchez-Cuadrado, S., Iglesias, A., & Berrios, O. (2020, December). An Accessible Evaluation Tool to Detect Easy-To-Read Barriers. In 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (pp. 55-60).	1) Medição da legibilidade do texto; 2) Detecção das dificuldades de legibilidade; 3) Indicação das dificuldades de legibilidade; 4) Sistemas de recomendação	Avaliação de acessibilidade linguística de textos	Uma ferramenta de avaliação acessível para detectar barreiras fáceis de ler baseada em regras (heurísticas) e processamento de linguagem natural (NLP)	Publicado em 2020 nos Proceedings do 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion
5 Antonelli, H. L., Sensiate, L., Watanabe, W. M., & de Mattois Fortes, R. P. (2019, October). Challenges of automatically evaluating rich internet applications accessibility. In Proceedings of the 37th ACM International Conference on the Design of Communication (pp. 1-6).	1) Técnicas de simulação de interações e reconhecimento de padrões podem ser incorporadas às ferramentas de avaliação atuais para avaliações mais precisas e redução de avaliações manuais. 2) Técnicas de IA podem ajudar a detectar componentes nas interfaces de RIAs e seus comportamentos e relacionamentos.	A avaliação automática de Rich Internet Applications (ou aplicações de internet rica)	O estudo apresenta propostas com algum detalhamento, mas ainda sem evidências específicas.	Publicado em 2019 nos Proceedings of the 37th ACM International Conference on the Design of Communication
6 Frug, S., & Bruce, T. (2019). Artificial Intelligence and Accessibility for Administrative Applications. In: Proceedings of the First Workshop on AI in the Administrative State.	1) Classificar e converter manualmente figuras em SVG e imagens de equações em MML; 2) Anotar imagens SVG com descrições de seu conteúdo; 3) Pesquisar fontes de dados legíveis por máquina representadas como imagens; 4) Aplicar técnicas de aprendizado de máquina para fornecer suporte para decisões de anotação e conversão humana.	Transformar informações matemáticas contidas em imagens em formatos mais acessíveis como MML e auxiliar no processo de anotação humana de outros tipos de imagens.	A abordagem identificou imediatamente 215 das 243 imagens matemáticas para conversão e identificou incorretamente apenas 35 das 875 imagens não matemáticas.	Publicado em 2019 nos Proceedings of the First Workshop on AI in the Administrative State (AIAS'19).

>> continúa

referência do trabalho	QP1: quais recursos e/ ou técnicas de IA/AM vem sendo testados na avaliação de acessibilidade web pela revisão de conformidade com o WCAG?	QP2: Que tipos de problema de avaliação de acessibilidade web se buscou resolver com a aplicação de IA e/ ou AM?	QP3: Os estudos apresentam evidências sobre vantagens da aplicação de IA / AM em relação a abordagens tradicionais de avaliação de acessibilidade por software?	QP4: Onde e quando os estudos encontrados foram publicados?
7 Morillo, P., Chicaiza-Herrera, D., & Vallejo-Huanga, D. (2019, July). System of recommendation and automatic correction of web accessibility using artificial intelligence. In International Conference on Applied Human Factors and Ergonomics (pp. 479-489). Springer, Cham.	Serviços de IA do Google (reconhecimento de imagem e descrição)	Obter uma descrição adequada dos itens multimídia que não estão rotulados corretamente.	Conformidade aumentou em média 2,57% após a correção em relação aos textos alternativos.	Publicado em 2019 nos Proceedings da International Conference on Applied Human Factors and Ergonomics (AHFE'19).

Tabela 2: estudos selecionados após etapa de leitura completa

Discussão

Inicialmente, percebe-se que os resultados indicam um número bastante baixo de estudos que efetivamente utilizam recursos de IA / AM na avaliação de acessibilidade a partir do recorte de pesquisa definido. Abaixo, são realizadas algumas análises para melhor compreender o cenário definido pelos achados. Os números fazem referências aos trabalhos na Tabela 2.

Em relação ao tipo de problema em que se aplicou técnicas de IA/AM:

- avaliação de descrição de imagens (1 e 7)
- detecção de tabelas e seus cabeçalhos (2)
- suporte aos desenvolvedores/avaliadores por informações sobre problemas via chatbot (3)
- avaliação da acessibilidade linguística (4)
- simulação de interações e reconhecimento de padrões para avaliações mais precisas e redução de avaliações manuais (5)
- conversão de imagens matemáticas em MML e anotação semântica (6)

Considerando-se a multiplicidade de questões passíveis de avaliação em relação à acessibilidade, percebe-se que há muitas áreas ainda a serem exploradas.

Em relação à forma em que foram publicados:

- três trabalhos finais de pós-graduação (dois de mestrado e um de doutorado, todos na área de tecnologia)
- quatro artigos em anais de eventos (com recortes diversos, predominando a área de tecnologia)

Dado que o tema é bastante recente, percebe-se que há quase metade entre as contribuições encontradas de trabalhos de pós-graduação, que provavelmente ainda não haviam sido publicados até a realização das buscas desta pesquisa em eventos ou periódicos.

Considerações Finais

A pesquisa em acessibilidade no contexto da Web trata de questões muito diversas, todas relacionadas com a qualidade da interação de pessoas com uma dada interface de um sítio eletrônico.

Vale destacar a importância de ter foco na pessoa com deficiência como impulsionador para mudanças no abismo histórico de falta de condições de acesso para esta significativa parte da população, bem como para o desenvolvimento de soluções para os casos específicos em cada categoria de deficiência.

Contudo, é necessário observar que o foco em projetos para todos, que minimizem restrições em vez de buscar soluções especializadas, tende a favorecer que sejam soluções acessíveis a mais pessoas passem a ser mais amplamente utilizadas em nossa sociedade, em relação às soluções desenvolvidas para perfis muito específicos de usuário.

O número vertiginosamente crescente de sites na internet estimula o desenvolvimento de novas soluções que permitam avaliar e se possível corrigir problemas de acessibilidade de forma mais efetiva e automática. Contudo, há um número ainda muito baixo de estudos que apresentam propostas concretas utilizando inteligência artificial e aprendizado de máquina. Entendemos que, soluções com técnicas de inteligência artificial e aprendizado de máquina podem ajudar o design a ter papel preponderante na transformação da sociedade, a partir da geração de projetos acessíveis para diferentes pessoas.

Como trabalhos futuros, pretende-se realizar novos estudos de revisão de literatura mais específicos, por exemplo para identificar questões mais problemáticas no design e avaliação de sítios eletrônicos para a acessibilidade, a fim de se verificar a viabilidade de propor ferramentas de apoio e recomendação que orientem resultados mais efetivos para a inclusão.

Referências

- Abou-Zahra, S., Brewer, J., & Cooper, M. (2018, April). Artificial Intelligence (AI) for Web Accessibility: Is Conformance Evaluation a Way Forward?. In *Proceedings of the 15th International Web for All Conference* (pp. 1-4).

- Aizpurua, A., Arrue, M., & Harper, S., Vigo, M. (2014). *Are Users the Gold Standard for Accessibility Evaluation?* W4A'14 Proceedings. <https://doi.org/10.1145/2596695.2596705>
- Antonelli, H. L., Sensiate, L., Watanabe, W. M., & de Mattos Fortes, R. P. (2019, October). Challenges of automatically evaluating rich internet applications accessibility. In *Proceedings of the 37th ACM International Conference on the Design of Communication* (pp. 1-6).
- BigDataCorp, & Movimento Web para Todos. (2022). *Acessibilidade na web brasileira | 4a edição.* <https://bigdatacorp.com.br/estudo-acessibilidade-na-web-brasileira-4a-edicao/>
- Brajnik, G., Yesilada, Y., & Harper, S. (2011). *The Expertise Effect on Web Accessibility.* Human–Computer Interaction, 2011, Volume 26. Taylor & Francis Group.
- Disability Rights Commission (DRC). (2004). *The web access and inclusion for disabled people: a formal investigation conducted by the Disability Rights Commission.* London: TSO. <https://disability-studies.leeds.ac.uk/wp-content/uploads/sites/40/library/DRC-Web-FI.pdf>
- Dogra, H. (2020). *A Framework for an automatic evaluation of image description based on an image accessibility guideline* (Master's thesis, OsloMet-storbyuniversitetet).
- Frug, S., & Bruce, T. (2019). Artificial Intelligence and Accessibility for Administrative Applications. In: *Proceedings of the First Workshop on AI in the Administrative State.*
- Haider, W. (2021). Automatically Detecting Table Headings In Web Pages With Machine Learning (Master's thesis, Middle East Technical University).
- Hertzum, M., Jacobsen, N., & Molich, R. (2002). *Usability inspections by groups of specialists: perceived agreement in spite of disparate observations.* Proceedings of CHI '02 (extended abstracts), pp. 662–663. ACM. <https://doi.org/10.1145/506443.506534>
- Hornbeak, K., & Frokjær, E. (2008). *A study of the evaluator effect in usability testing.* Human-Computer Interaction, 23(3):251 – 277. <https://doi.org/10.1080/07370020802278205>
- IBGE - Instituto Brasileiro de Geografia e Estatística. (2021). *Pesquisa nacional de saúde: 2019: ciclos de vida.* <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101846.pdf>
- International Organization for Standardization (ISO). (2008). *Ergonomics of human-system interaction — Part 171: Guidance on software accessibility* (ISO 9241-171). ISO. <https://www.iso.org/standard/39080.html>
- Jacobsen, N., Hertzum, M., & John, B. (1998). *The evaluator effect in usability tests.* Proceedings of CHI '98. ACM. <https://doi.org/10.1145/286498.286737>
- Kavaklıoglu, E. AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference? (2020, May 27). <https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>
- Kitchenham, B. A., Budgen, D., & Brereton, O. P. (2010, April). *The value of mapping studies—A participant-observer case study.* In 14th international conference on evaluation and assessment in software engineering (ease) (pp. 1-9).
- Lopes, R., Gomes, D., & Carriço, L. (2010). Web not for all: a large scale study of web accessibility. Proceedings of the W4A'10. <https://doi.org/10.1145/1805986.1806001>
- Meirelles, F. S. *Panorama do Uso de TI no Brasil - 2022.* (2022, May 26). Portal FGV. <https://portal.fgv.br/artigos/panorama-uso-ti-brasil-2022>

- Morato, J., Campillo, A., Sanchez-Cuadrado, S., Iglesias, A., & Berrios, O. (2020, December). An Accessible Evaluation Tool to Detect Easy-To-Read Barriers. In 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (pp. 55-60).
- Morillo, P., Chicaiza-Herrera, D., & Vallejo-Huanga, D. (2019, July). System of recommendation and automatic correction of web accessibility using artificial intelligence. In *International Conference on Applied Human Factors and Ergonomics* (pp. 479-489). Springer, Cham.
- Organização das Nações Unidas. (2008). *Convenção sobre os Direitos das Pessoas com Deficiência*. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/convention-on-the-rights-of-persons-with-disabilities-2.html>
- Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008, June). Systematic mapping studies in software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering* (EASE) 12 (pp. 1-10).
- Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and software technology*, 64, 1-18.
- Power, C., Freire, A. P., Petrie, H., & Swallow, D. *Guidelines are only half of the story: accessibility problems encountered by blind users on the web*. CHI '12 Proceedings. <https://doi.org/10.1145/2207676.2207736>
- Rodriguez, G. (2020). *Improving Web Accessibility Through Suggestions Using Serverless Architecture* (Doctoral dissertation, California State University, Northridge).
- Russel, S.J., & Norvig, P. *Artificial Intelligence: A Modern Approach*. (2016). New Jersey: Prentice Hall.
- Souza, Edson Rufino de. (2005). *Anais do ATIID 2005*. São Paulo: Prodam. http://www.prodam.sp.gov.br/multimidia/midia/cd_atiid/conteudo/atiid2005/mr2/02/avaliacaowebsite.pdf
- Souza, Edson Rufino de, & Mont'Alvão, Claudia. (2012). *Web accessibility: Evaluation of a website with different semi-automatic evaluation tools*. Work, 41, 1567–1571. <https://doi.org/10.3233/WOR-2012-0354-1567>
- Souza, Edson Rufino de. (2015). *Avaliação da metodologia de conformidade com o Web Content Accessibility Guidelines (WCAG-EM 1.0)* [Doutorado em Design, Pontifícia Universidade Católica do Rio de Janeiro]. https://www.dbd.puc-rio.br/pergamon/tesesabertas/1113325_2015_completo.pdf
- Souza, Edson Rufino de. (2016). Acessibilidade web: diferentes definições e sua relação com o design universal. *Diálogo com a Economia Criativa*, v. 1, p. 13, 2016. <http://dx.doi.org/10.22398/2525-2828.1113-28>
- Waller, S. & Clarkson, P. J. Tools for Inclusive Design. In: C. Stephanidis (Ed.). *Universal Access Handbook*. CRC Press.
- WebAIM. (2022, March 31). *The WebAIM Million - The 2022 report on the accessibility of the top 1,000,000 home pages*. <https://webaim.org/projects/million/>
- World Wide Web Consortium. *Web content accessibility guidelines 2.0*. 2008. <http://www.w3.org/TR/WCAG20/>.

World Wide Web Consortium. *Web content accessibility guidelines 2.1.* 2018. <http://www.w3.org/TR/WCAG21/>.

World Wide Web Consortium. *Introduction to Web Accessibility.* 2022. <https://www.w3.org/WAI/fundamentals/accessibility-intro/>.

Yesilada, Y., Brajnik, G., & Harper, S. (2009). *How Much Does Expertise Matter? A Barrier Walkthrough Study with Experts and Non-Experts. Assets' 2009 Proceedings.* <https://doi.org/10.1145/1639642.1639678>

Abstract: Web accessibility is essential for the full participation of people with or without disabilities in society. However, most sites do not meet the minimum accessibility requirements. Accessibility assessment is one of the steps to make websites more accessible. We present related challenges and discuss the current possibilities of using artificial intelligence and machine learning to improve the accessibility of websites on the Internet.

Keywords: accessibility - artificial intelligence - machine learning

Resumen: La accesibilidad web es fundamental para la plena participación de las personas con o sin discapacidad en la sociedad. Sin embargo, la mayoría de los sitios no cumplen con los requisitos mínimos de accesibilidad. La evaluación de la accesibilidad es uno de los pasos para hacer más accesibles los sitios web. Presentamos desafíos relacionados y discutimos las posibilidades actuales de utilizar la inteligencia artificial y el aprendizaje automático para mejorar la accesibilidad de los sitios web en Internet.

Palabras clave: accesibilidad - inteligencia artificial - aprendizaje automático

[Las traducciones de los abstracts fueron supervisadas por su autor]
