

## معیاری برای ارزیابی دسترس پذیری وب\*

\*بم بنگ پارمانتو و شیاومینگ زنگ

(گروه مدیریت اطلاعات سلامت، دانشکده علوم سلامتی و

بازتوانی، دانشگاه پیتسبورگ، پیتسبورگ، پنسیلوانیا)

\*ترجمه سمیه سادات آخشیک

کارشناس ارشد کتابداری و اطلاع رسانی

somakhshik@gmail.com

پژوهشگاه علوم انسانی و مطالعات فرهنگی  
رتال جامع علوم انسانی

پذیری وب در مقیاس کلان با استفاده از این معیار آورده شده است. این مثال ها، مقایسه ای از دسترس پذیری وب، از وب سایت های مجلات سطح بالای پزشکی و مطالعه طولی یک وب سایت را در طی زمان، دربر می گیرند. اعتبار این معیار با استفاده از تعداد زیادی وب سایت با سطوح مختلف مطابقت (رتبه بندی نمودن دسته ها) با استاندارد راهنمای دسترس پذیری محتوای وب مورد آزمایش قرار گرفت.

این معیار، که از طرح توزین ساده از پیش تعیین شده استفاده می کند، در جداسازی وب سایت ها به دسته های رتبه بندی مختلف به خوبی با الگوریتم پیچیده تر یادگیری ماشینی C5.0 در جداسازی وب سایت ها به دسته های رتبه بندی مختلف قابل قیاس است.

معیار جدیدی برای اندازه گیری کمی دسترس پذیری محتوای وب، برای افرادی که معلولیت دارند پیشنهاد شده است. معیار مبتنی است بر راهنما های دسترس پذیری محتوای وب (WCAG)<sup>1</sup> که یک استاندارد پذیرفته شده بین المللی است و می تواند با استفاده از برنامه های کامپیوتری به طور خودکار آزمایش شود. مشکلات مربوط به ارزیابی دسترس پذیری محتوای وب و نیاز به معیار دسترس پذیری وب مورد بحث قرار گرفته است. معیار پیشنهاد شده، بر آن است تا بر کاستی های روش های اندازه گیری مورد استفاده در مطالعات دسترس پذیری وب، غلبه نماید. معیار پیشنهادی، با الزامات یک ابزار سنجش در تحقیقات علمی مطابقت دارد. مثال هایی از ارزیابی های دسترس

\* Parmanto, Bambang; Zeng, Xiaoming. Metric for web accessibility evaluation. Journal of the society for information science and technology, 56(13):1394-1404,2005.

## مقدمه

اهمیت اندازه گیری ویژگی های ارقام شناخته شده از نظر کمی، در پیشبرد موقعیت علم در هر رشته ای، تعیین کننده است. وب، به عنوان یکی از جالب توجه ترین ابزارهای جدید تحقیق، معیارهای زیادی را برای کمک به بررسی ای علمی ایجاد نموده است (Dhyani, Wee keong & Bhowmick, 2002).

در این مقاله، ما معیار جدیدی را برای اندازه گیری دسترس پذیری محتوای وب برای افراد معلول پیشنهاد می کنیم. اندازه گیری دسترس پذیری وب با معیارهای دقیق و کمی، به دلایل زیادی مهم می باشد.

نخست اینکه، می تواند به طور کلی آگاهی ما را از وب افزایش دهد. همچنین به ما این امکان را می دهد که وضعیت فعلی دسترس پذیری وب را اندازه گیری نماییم و دسترس پذیری وب سایت های مختلف و نیز دسترس پذیری یک وب سایت خاص را در زمان های مختلف مقایسه نماییم. روش اندازه گیری متوالی عددی به روش اندازه گیری دو مرحله ای کنونی دسترس پذیری، ترجیح خواهد داشت. یک مقیاس پیوسته نه تنها اندازه گیری دقیق تر دسترس پذیری را ممکن می سازد، بلکه با تحلیل های آماری پیشرفته تر برای ارزیابی وب سایت های بزرگ در مقیاس کلان نیز سازگاری دارد.

شیوه کنونی ارزیابی دسترس پذیری وب، از یک روش دو مرحله ای مبنی بر پذیرش مطلق راهنما های استاندارد استفاده می کند که به عنوان راهنما های دسترس پذیری محتوای وب (WCAG) 1.0 شناخته شده و توسط کنسرسیوم تور جهان گستر (W3C) تهیه شده است. قابلیت دسترسی یا دسترس ناپذیری یک وب سایت با ارزیابی آن در مقایسه با نقاط بازرسی ارائه شده توسط WCAG تعیین می گردد. WCAG شامل 14 راهنما عبارت بندی شده گسترده می باشد که به 91 نقطه بازرسی ویژه تبدیل تفسیر شده اند که توضیح می دهند چگونه راهنما ها باید در مورد سناریو های خاص ایجاد محتوی به کار روند.

این نقاط بازرسی در سه سطح اولویت سازمان دهی شده اند: اولویت اول شامل 29 نقطه بازرسی می باشد که حتماً باید برآورده شوند؛ اولویت دوم شامل 40 نقطه بازرسی می باشد که بایستی برآورده شوند و اولویت سوم شامل 22 نقطه بازرسی می باشد که ممکن است برآورده شوند. با در نظر گرفتن تعداد نقاط بازرسی که وب

سایت حتماً باید مطابق آنها باشد تا دسترس پذیر تلقی گردد، تعجب آور نیست که نتایج مطالعات دسترس پذیری، اغلب وب سایت ها را غیر قابل دسترس یافت. حتی تطابق با دسترس پذیری مقدماتی در اولویت اول، دشوار خواهد بود. در این روش اندازه گیری دو بخشی هر گونه تخطی از 29 نقطه بازرسی موجود در اولویت اول، از قبیل فراموش کردن تخصیص متن جایگزین برای یکی از تصاویر روی وب سایت، آن را دسترس ناپذیر خواهد ساخت. همچنین این نوع روش اندازه گیری دو بخشی منجر به بی دقتی در تعیین دسترس پذیری می گردد. یعنی اکثر وب سایت هایی که ادعای دسترس پذیری کامل دارند، در حقیقت از راهنما هایی که انتظار می رود با آنها مطابقت دارند تخطی می کنند. مطالعه ما نشان داد که تنها 8/81% از وب سایت هایی که ادعا می کنند در سطح بسیار مطلوبی هستند (با نقاط بازرسی اولویتها ی 1،2 و 3 از WCAG مطابقت دارند) بدرستی مطابق با هر سه اولویت می باشند ( به جدول 1 نگاه کنید).

همچنین روش اندازه گیری متداول دسترس پذیری، اندازه و پیچیدگی یک وب سایت را به حساب نمی آورد. یک وب سایت بزرگ با صدها یا هزاران صفحه وب، نسبت به وب سایتی ساده که فقط شامل تعداد اندکی وب است، احتمال بیشتری برای تخلف از نقاط بازرسی خواهد داشت. معیار اندازه گیری دسترس پذیری که اندازه و پیچیدگی را به حساب می آورد، امکان مقایسه منصفانه ای را میان وب سایت ها یا مجموعه ای از وب سایت ها را خواهد داد.

سالیوان و ماتسن<sup>3</sup> (2002) اولین کسانی بودند که ایده روش اندازه گیری مداوم دسترس پذیری را پیشنهاد نمودند. یعنی اندازه گیری دسترس پذیری بر حسب «درجه ها» به جای اندازه گیری دو وجهی دسترسی - عدم دسترسی. ولی آنها جزئیات محاسبه اندازه گیری مداوم را مطرح نکردند. در عوض، وب سایت ها را در 4 درجه دسترس پذیری رتبه بندی نمودند؛ دسترس پذیری عالی، دسترس پذیری زیاد، دسترس پذیری نسبی و غیر دسترس پذیر. یک معیار عددی با مقادیر پیوسته، قدرت تمایز بهتری را فراهم می نماید و رویکردی علمی برای مباحث دسترس پذیری وب به وجود می آورد.

در این مقاله، ما معیار جدیدی را برای اندازه گیری کمی دسترس پذیری وب پیشنهاد می کنیم.

این معیار نقطه شروعی است که از راهنماهای WCAG استفاده می‌کند. به طور دقیق‌تر، این معیار بر اساس نقاط بازرسی WCAG می‌باشد که می‌تواند با استفاده از برنامه‌های کامپیوتری به طور خودکار آزمایش شود. این معیار به عنوان برآوردی از دسترس پذیری در نظر گرفته شده است، حال آنکه مقیاس‌های واقعی دسترس پذیری (خودکار یا طور دیگر) نیازمند کنترل اضافی دستی و داوری انسانی هستند.

### پیشینه و کارهای مرتبط

#### راهنماهای دسترس پذیری وب

راهنماهای متعددی برای کمک به طراحان وب در دسترس پذیر ساختن وب سایت‌ها برای افراد معلول تهیه شده است. در دهه ۱۹۹۰، اطلاعات دسترس پذیری وب از طریق سازمان‌هایی مانند مرکز تحقیق و توسعه تریس<sup>۴</sup> در دانشگاه ویسکانسین و شرکت‌هایی از قبیل IBM فراهم شد. یکی از اولین استانداردهای طراحی محتوی وب برای افراد معلول، در شهر سن خوزه کالیفرنیا تهیه شد (Paciello, 2000). در ۱۹۹۷، استانداردهای استرالیایی برای طراحی وب دسترس پذیر، برای نویسندگان صفحه وب فراهم شد (Australian Human Rights & Equal Opportunity Commission, 1997).

در همان سال، W3C، پایه‌گذار پیشگامان دسترس پذیری وب (WAI) گردید. WCAG 1.0 در سال ۱۹۹۹، WAI را به عنوان توصیه‌نهایی به انتشار رسانید (WAI, 1999). دو ویژگی اصلی وجود دارد که به عنوان راهنماهای بهنجار برای طراحی دسترس پذیری محتوای وب بکار می‌روند: WCAG و استاندارد‌های دسترس پذیری تکنولوژی اطلاعات و الکترونیکی هیأت دسترس‌یابی ایالات متحده (فصل ۵۰۸ راهنماها).

WCAG، یک مجموعه خصوصیات بین‌المللی است که از طریق یک توافق صنعتی داوطلبانه تهیه شده است.

راهنماهای فصل ۵۰۸ در امتداد فرآیند قانونی سازی ایالات متحده که لازمه فصل ۵۰۸ از اصلاحیه قانون توانبخشی است، در دسامبر ۲۰۰۰ اعلام شد. هر دو مجموعه، چک لیستی را پیشنهاد می‌کنند که سازندگان وب باید با در نظر گرفتن دسترس پذیری محتوا برای افراد معلول، از آن پیروی نمایند. این دو مبنای، تا حد زیادی همپوشانی دارند،

فقط سه نقطه بازرسی که در فصل ۵۰۸ تعیین شده اند، در WCAG ذکر نشده‌اند. WCAG، نقاط بازرسی جامع‌تری نسبت به فصل ۵۰۸ دارد و برای هر نقطه بازرسی، سطح اولی‌تری را تعیین می‌کند که شدت نقض مقررات مربوطه را نیز منعکس می‌کند. WCAG شامل ۱۴ دستورالعمل می‌باشد که به ۹۱ نقطه بازرسی ویژه تبدیل شده‌اند که بیانگر چگونگی بکارگیری دستورالعمل‌ها در سناریوهای خاص ساخت محتوا می‌باشند. این نقطه بازرسی‌ها در سه سطح از اولویت سازمان‌دهی شده‌اند: اولویت اول شامل ۲۹ نقطه بازرسی می‌باشد که حتماً باید برآورده شوند؛ اولویت دوم شامل ۴۰ نقطه بازرسی می‌باشد که بایستی برآورده شوند؛ و اولویت سوم شامل ۲۲ نقطه بازرسی می‌باشد که می‌توانند برآورده شوند.

پیشگامان دسترس پذیری وب<sup>۵</sup>، برای ارتقاء بیشتر دسترس پذیری وب، نشانه‌های تطابق WCAG را معرفی نموده است. تهیه‌کنندگان محتوی می‌توانند از این نشانه‌ها در وب سایت‌های خود برای نشان دادن میزان تطابق با سطوح ویژه WCAG استفاده نمایند. پیشگامان دسترس پذیری وب، انتظار دارد که استفاده از این نشانه‌ها، در وب سایت‌های همخوان، به افزایش آگاهی در مباحث دسترس پذیری کمک می‌نماید. سطوح تطابق دارای تعاریف زیر هستند:

- تطابق سطح A: تمامی نقاط بازرسی اولویت ۱، حاصل شده‌اند.
- تطابق سطح AA: تمامی نقاط بازرسی اولویت ۱ و ۲ حاصل شده‌اند.
- تطابق سطح AAA: تمامی نقاط بازرسی اولویت ۱، ۲ و ۳ حاصل شده‌اند.

#### نیاز به معیاری برای دسترس پذیری

زمانی که WCAG توسط W3C پذیرفته شد و فصل ۵۰۸ قانون توانبخشی بصورت قانون در آمد، مطالعات متعددی درمورد دسترس پذیری دسته‌های گوناگون وب سایت‌ها انجام شده است.

این مطالعات از WCAG به عنوان مبنای اندازه‌گیری دسترس پذیری استفاده نمودند و ابزار سنجش خودکار، را Bobby (watchfirecrsop, 2004) برای ارزیابی بکار بردند.

چنین مطالعاتی معمولاً تصویری تیره از وضعیت دسترس پذیری وب نمایش می‌دهند.

مطالعه ای که اخیراً از دسترس پذیری ۳۰ وب سایت بسیار معروف فرانسوی انجام شد، مشخص نمود که هیچ یک از آنها با سطح A مطابقت ندارند (Research Institute for Networks and Communications Engineering [RINCE], 2003)

مطالعه مشابهی که در ایرلند انجام شد مشخص نمود که دست کم ۹۴٪ از ۱۵۹ وب سایت آزمایش شده، با کمترین استاندارد دسترس پذیری (سطح A) مطابقت نداشتند و حتی یک سایت هم با راهنما دسترس پذیری حرفه ای عملی از سطوح A و AAA مطابقت نداشت (McMullin, 2002).

نتایج مطالعات پیشین اغلب سردرگم کننده و متعارض می باشند. مطالعه ای که در مورد دسترس پذیری وب سایت های فدرال امریکا انجام شد، مشخص نمود که تنها ۱۳/۵٪ از ۱۴۸ سایت، فاقد خطا هستند (Stower, 2002) که مبین آن است که این وب سایت ها را می توان به عنوان سطح AAA یا « تأیید شده ی Bobby» در نظر گرفت. این مطالعه بسیار مشهور شده است (Deaukantas, 2002; Emery, 2002) که بخشی از آن به این علت است که تصور می شد تمامی وب سایت های فدرال امریکا تا ۲۵ ژوئن ۲۰۰۱ با قانون توانبخشی سال ۱۹۷۳ مطابقت داشته باشند.

در مطالعه ای که پیش از این توسط محققان دانشگاه براون انجام شد معلوم گردید که ۳۷٪ از وب سایت های دولتی امریکا قابل دسترسی می باشند (West, 2001).

مطالعه دیگری مشخص کرد که تنها ۱٪ از وب سایت های دولتی فدرال امریکا از نوع Bobby می باشند (Jackson-Sanborn, Odess-Hamish, & Warren, 2002) که در آن جا بعنوان مطابقت با اولویت ۱ (سطح A) اما بدون کنترل کاربر تعریف شده بود. تمامی این مطالعات از Bobby که یک ابزار خودکار ارزیابی دسترس پذیری است و از مقیاس مطلق دسترس پذیری استفاده نموده اند. نرخ پایین دسترس پذیری در بین وب سایت های دولتی داستان خوبی برای رسانه ها خواهد شد، اما به قدر کافی برای اهداف علمی یا سیاسی آموزنده نخواهد بود.

بحث ما این است که ابهام نتایج مغایر، ناشی از مشکلات معیار های به کار رفته در خود مطالعات می باشد. روش کنونی ارزیابی دسترس پذیری وب سایت، متکی به رتبه بندی ساده ای است مبنی بر تطابق با مجموعه نقاط بازرسی اولویت بندی است که در WCAG ارائه شده اند. نظام رتبه

بندی فعلی و به اصطلاح روش اندازه گیری تأیید شده Bobby، منعکس کننده معیار قطعی دسترس پذیری وب می باشند. یا اینکه سایت با تمام نقاط بازرسی مطابقت دارد، یا اینکه غیر قابل دسترسی محسوب می گردد.

به منظور نشان دادن مشکلات کنونی روش اندازه گیری دو وجهی مطلق دسترس پذیری، یک ارزیابی بر روی نمونه وسیعی از وب سایت هایی که خود را دسترس پذیر در نظر گرفته اند، انجام دادیم. ۴۴۹، ۳۷۴ و ۳۱۸ وب سایت را که خود را به ترتیب در سطوح AA، A و AAA رتبه بندی کرده بودند، برای ۱۱۴۱ وب سایت انتخاب نمودیم (برای اطلاعات بیشتر در مورد وب سایت های نمونه، به قسمت استاندارد طلایی بنگرید). سپس بیش از ۳۷۷ وب سایت را اضافه کردیم که با اولویت ۱ مغایرت داشتند و دسترس ناپذیر محسوب می شوند.

ما دسترس پذیری تمامی ۱۵۱۸ وب سایت را به منظور بررسی میزان تطابق آنها با هر یک از سطوح اولویت WCAG ارزیابی نمودیم. فقط آن دسته از نقاط بازرسی را که می توانند بصورت خودکار یا استفاده از برنامه های کامپیوتری ارزیابی شوند استفاده نمودیم. نتایج این ارزیابی در جدول ۱ ارائه شده است. حیرت آور است که حتی در میان وب سایت هایی که خود را دارای سطح تطابق AAA تلقی کردند، تنها ۸/۸٪ آنها واقعا AAA می باشند. تعدادی از مطالعات گذشته از معیار های AAA یا مصوبه Bobby به عنوان معیاری برای دسترس پذیری استفاده نمودند (Jackson-Sanborn et al, 2002; Stower, 2002). درصد وب سایت های مطابق با معیار AAA که خود را وب سایت های AA و A معرفی نموده بودند به طور قابل توجهی پایین است (به ترتیب، ۴/۲۸٪ و ۱/۱۱٪) و برای وب سایت هایی که بصورت تصادفی انتخاب شده بودند، تطابق با استاندارد AAA به صفر می رسد.

اگر بررسی تمامی ۹۱ نقطه بازرسی بصورت دستی انجام می شد و اگر تمامی صفحات وب سایت ها ارزیابی شده بودند، نسبت به بررسی ۲۵ نقطه بازرسی تنها با استفاده از روش تمام خودکار و فقط روی صفحه اصلی وب سایت، نتایج بدتر می شد.

این جدول، می تواند نتایج مطالعات پیشین



در بین ۱۵۹ وب سایت ایرلندی را شرح دهد (McMullin, 2002) که نشان دادند میزان عدم تطابق با سطح A، ۹۳/۷٪ و میزان عدم تطابق با AA و AAA، ۱۰۰٪ می باشد.

مطالعات مربوط به وب سایت های ایرلندی و فرانسوی (۳۰ وب سایت پر بازدید با ۱۰۰٪ میزان عدم تطابق با سطح A)، با نتایجی که از انتخاب تصادفی وب سایت ها بدست آوردیم، سازگار می باشد.

نتایج بسیار ناامید کننده ما، نشان دهنده ضعف مقیاس قطعی دسترس پذیری است که در مطالعات به کار رفته است.

Accessibility Level	Total	Failed	Passed
AAA	277	277	0
AA	172	172	0
A	172	172	0

#### جدول ۱. درصد وب سایت هایی با تخلف از اولویت های مبنی بر ۲۵ نقطه بازرسی

از آنجا که نقض تنها یک نقطه بازرسی در یک اولویت، وب سایت را دسترس ناپذیر می نماید، تنها درصد کمی از وب سایت ها می توانند دسترس پذیر تلقی شوند. این چنین نتایجی نمی توانند کمک چندانی به روشن نمودن وضعیت دسترس پذیری وب باشند. معیار متفاوت و بهتری برای اکتشافات علمی و نیز برای فرمول بندی های سیاسی مورد نیاز است.

#### نیاز به ارزیابی خودکار

تعداد صفحات منحصر به فرد وب، در ژوئیه ۲۰۰۰، ۲/۱ بیلیون صفحه برآورد شده است، که با نرخ رشد ۷/۵ میلیون صفحه در هر روز رو به افزایش می باشند (Murry & Moore, 2000). تعداد کل صفحات وب کاملاً پنهان، و پایگاه های داده متصل به وب به ۵۵۰ بیلیون مدرک وب نامرئی تخمین زده می شود (Bergman, 2001). وجه مشخصه وب، تنها گستردگی صرف آن نیست بلکه سیال بودن آن نیز می باشد؛ وب سایت ها به طور مداوم تغییر می کنند. چنانکه پروژه آرشیو اینترنتی (Weiss, 2003) برآورد نموده است، متوسط طول عمر یک صفحه وب، اکنون ۱۰۰ روز است. مطالعه ای که (نتیجه آن) در مجله ساینس منتشر

شد، نشان داد که میزان رواج اسنادهای غیر فعالی که در مجلات ممتاز به وب می شود، پس از ۱۵ ماه ۱۰٪ می باشد (Dellavale, et al, 2003). با توجه به ماهیت وب، امتیاز بندی و ارزیابی خودکار، نسبت به امتیاز بندی دستی، برتر و مؤثرتر می باشد. ارزیابی خودکار دسترسی وب، مزیت های بسیاری در مقایسه با ارزیابی غیر خودکار دارد، مانند هزینه انجام ارزیابی، زمان مورد نیاز برای انجام ارزیابی، یکدستی فزاینده در دسترس پذیری آشکار، کاهش نیاز به مهارت در دسترس پذیری و امکان گنجاندن ارزیابی دسترسی در فرآیند ساخت وب. در مورد قابلیت استفاده ارزیابی خودکار، بحث های مشابهی انجام شده است (Ivory & Hearst, 2001). حتی یک بحث حتی استدلال متقاعد کننده تری در خصوص ارزشیابی خودکار دسترسی به وب حتی یک ادعای محکم تر وجود دارد: دستورالعملی با نقاط بازرسی مفصل وجود دارد که دارای پذیرش بین المللی است. امتیاز بندی خودکار این امکان را می دهد که ارزیابی روی تعداد زیادی وب سایت و در زمان اندک و با حداقل هزینه انجام پذیرد.

#### ویژگی های یک معیار دسترس پذیری سودمند

##### وبی

به منظور غلبه بر نواقص معیار قطعی کنونی، معیاری برای دسترس پذیری پیشنهاد می کنیم که چندین شرط لازم را دارا باشد. اول اینکه دسترس پذیری باید با نمراتی سنجیده شود که مقادیر پیوسته ای از کاملاً دسترس پذیر تا کاملاً غیر قابل دسترس را دربر گیرد. امتیازات کمی و عددی ارزیابی تغییرات دسترس پذیری وب را در طول زمان و نیز مقایسه ی بین وب سایت ها یا بین گروه های وب سایت را امکان پذیر می سازد. به جای مقیاس مطلق دسترس پذیری که وب سایت ها را فقط از نظر دسترسی یا عدم دسترسی طبقه بندی می نماید، ارزیابی با استفاده از این معیار قادر خواهد بود به یک پرسش اساسی علمی پاسخ گوید: دسترس پذیری کمتر یا بیشتر در مقایسه با چه چیز؟ (Tuft, 1997). دوم اینکه معیار و دامنه مقادیر، باید قدرت تمایز زیادی را که فراتر از تعیین صرف دسترس پذیری یا عدم دسترس پذیری است دارا باشد. معیاری با قدرت تمایز خوب، امکان بررسی میزان تغییرات دسترس پذیری وب را در طی زمان، یا بروز تفاوت قابل ملاحظه در دسترس پذیری بین وب سایت های

تحت بررسی را فراهم می آورد. ارزیابی دسترس پذیری با استفاده از این معیار، قادر است به این سئوال اساسی علمی پاسخ دهد: با چه نرخ؟ (Tuft, 1997).

سوم، این معیار بایستی آنقدر منصفانه باشد که اندازه و پیچیدگی های وب سایت ها را به حساب آورده و خود را با آن ها هماهنگ سازد. دامنه وب سایت ها ممکن است از یک صفحه خانگی تا سایت های بزرگ شرکت ها را که شامل هزاران صفحه می باشند، دربر بگیرد. معیاری که اندازه و پیچیدگی را به حساب می آورد، امکان مقایسه منصفانه ای را بین وب سایت ها با اندازه های گوناگون فراهم می آورد.

چهارم، این معیار باید برای اجرای مطالعات دسترس پذیری وب در مقیاس وسیع، مقیاس پذیر باشد. ارزیابی های دسترس پذیری در مقیاس وسیع، نیازمند معیاری است که از آمارهای تجمعی و رده دوم مانند انحراف معیار پشتیبانی کند. برای یک مطالعه در مقیاس وسیع، برجسته ترین امر، کارایی است.

بالاخره، روش اندازه گیری، باید بهنجار باشد، به این معنی که باید برگرفته از راهنما های استاندارد دسترس پذیری وب از قبیل WCAG یا فصل ۵۰۸ باشد.

معیار پیشنهادی برای کار با یک روش ارزیابی خودکار دسترس پذیری طراحی شده است. اگرچه این معیار، دسترس پذیری وب را از روی پراکسی نشان می دهد و نه از روی تجربه کاربر، اما عملی است و توانایی های زیادی دارد. یکی از مهمترین نقاط قوت روش امتیاز بندی خودکار، عینیت آن است؛ امکان مقایسه عینی بین سایت ها، دسته ها، و زمان های مختلف را می دهد. این روش همچنین امکان ارزیابی های گسترده را از انبوه وب سایت ها می دهد. ارزیابی تطابق تمامی نقاط بازرسی WCAG به طور بالقوه نیازمند این است که هر صفحه از وب سایت به طور مفصل با هر نقطه بازرسی مقایسه شود. حال ارزیابی ۱۰۰ وب سایت را که شامل ۱۰۰۰ صفحه وب در هر سایت باشد را تصور کنید. ارزیابی دسترس پذیری وب در مقیاس وسیع با استفاده از روش ارزیابی دستی، هم زمانبر و هم به طور بازدارنده ای گران<sup>۷</sup> است.

همانطور که مک مالین<sup>۸</sup> (۲۰۰۲) استدلال نموده، ارجح این است که داده های مشخص و جامع و

مرتبط با دسترس پذیری وب به میزان فراوان در اختیار باشد، حتی اگر که این داده ها کامل نباشند.

## معیار جدید دسترس پذیری: مانع دسترس

### پذیری وب

یکی از نتایجی که از بررسی پیشینه ها می توانیم بگیریم این است که روش های ارزیابی پذیرفته شده کنونی برای

دسترس پذیری وب، دو ضعف عمده دارند. نخست اینکه اغلب آنها تنها تعداد از پیش تعیین شده ای از خطاهای موجود دسترس پذیری وب را در یک صفحه وب در نظر می گیرند. تنها شمارش تعداد خطاها، بدون در نظر گرفتن تعداد خطاهای بالقوه، مانند تعداد عناصر تصویری در هنگام بررسی تصاویر متنی غیر جایگزین، به سود صفحاتی است که طراحی ساده ای دارند و ممکن است تلاش طراح وب در وب سایت های پیچیده، ناچیز انگاشته شود. دوم اینکه اغلب ارزیابی های دسترس پذیری وب مطالعات را به صورت یک نمره واحد سنجش واحد که مجموعه موانع دسترس پذیری یک صفحه وب یا وب سایت را نشان می دهد، ارائه نمی کنند. در عوض، نتایج اغلب به صورت طبقه بندی شده و بر مبنای نقاط بازرسی، راهنما ها یا اولویت های WCAG ارائه می شوند. اگرچه ارائه نمودن نتایج می تواند طرحی اجمالی از توزیع دسترس پذیری وب در بین وب سایت های مختلف فراهم نماید، اما استفاده صرف از این روش اندازه گیری طبقه ای، برای مقایسه دو صفحه وب، دشوار است.

همانطور که یکی از تحقیقات نشان داده است (Ivory & Hearst, 2002) این ضعف ها می تواند توضیح دهد که از چه جهت خطای دسترس پذیری Bobby زمانی افزایش می یابد که صفحات بهتر طراحی شده اند؛ بدلیل اینکه WCAG و فصل ۵۰۸ تا حد زیادی همپوشانی دارند و با توجه به این که WCAG استاندارد جامع تر و پذیرفته شده ای در سطح بین المللی است، ما WCAG را به عنوان مبنایی برای معیار دسترس پذیری ساخت خود به کار بردیم. تعداد خطاهای هر نقطه بازرسی قسمتی از روش امتیاز بندی ما را تشکیل می دهد که نمره مانع دسترس پذیری وب<sup>۹</sup> (WAB) نامیده می شود. به عنوان مثال، یک صفحه وب با خطاهای کمتر در نقاط بررسی نقطه بازرسی دسترس پذیری، مانند عدم تهیه توصیف جایگزین برای یک مقوله

تصویری، برای افراد معلول دارای موانع کمتری به حساب خواهد آمد و نمره WAB پایین‌تری دریافت خواهد داشت. از آنجا که علاقه ما بیشتر به ارزیابی خودکار سطح دسترسی یک وب سایت است، آن دسته از نقاط بازرسی که نیازمند بررسی دستی می‌باشند، در محاسبه امتیاز WAB در نظر گرفته نمی‌شوند. به عنوان مثال، پیروی از قاعده « اگر برای رساندن اطلاعات از رنگ استفاده می‌کنید، مطمئن شوید که اطلاعات به روش دیگری هم ارائه شده است. » تا زمانی که یک بررسی دستی انجام نگیرد، نمی‌تواند تأیید گردد. برای فهرستی از قواعد دسترسی پذیری وب که نیازمند بررسی دستی می‌باشد، به منبع WCAG نگاه کنید (WAI, 1999). همانطور که در قسمت پیشینه گفته شد، WCAG بر اساس تأثیر آن بر دسترسی پذیری افراد معلول، یک سطح اولویت سه ماده ای به هر نقطه بازرسی، ضمیمه می‌نماید. در وزن دادن به محاسبات نمره WAB، از سطوح اولویت به ترتیب عکس استفاده نمودیم. به این معنی که خطاهای اولویت ۱، سه برابر خطاهای اولویت ۳ می‌باشند، چرا که افراد معلول برای دسترسی به صفحات وب با اولویت ۱ دشواری بیشتری دارند. به هر حال، استفاده صرف از تعداد خطاهای نقاط بازرسی دسترسی پذیری وب ممکن است سبب سوگیری در نتایج روش اندازه‌گیری شود. به عنوان مثال، یک صفحه وب، با « پنج خطا در تصویر بدون متن جایگزین » ممکن است ۵۰۰ مورد تصویر جاسازی شده در صفحه را داشته باشد و صفحه وبی با « یک خطای در تصویر بدون متن جایگزین » ممکن است تنها یک مورد تصویر در صفحه داشته باشد. تهیه کننده‌ی صفحه اول ممکن است از قبل توجه و تلاش زیادی را برای تطابق با خصوصیات دسترسی پذیری وب به عمل آورده باشد، حال آنکه تهیه کننده‌ی صفحه دوم ممکن است که به کلی از دسترسی پذیری وب نا آگاه باشد. بنابراین، تعداد خطاهای واقعی یک نقطه بازرسی باید در برابر تعداد خطاهای بالقوه یک نقطه بازرسی، تعدیل شود. در مثال اخیر، خطاهای حقیقی، موارد تصویری بدون متن جایگزین می‌باشند و خطاهای بالقوه، شامل تمامی موارد تصویری روی صفحه هستند. نمره میانگین WAB همه صفحات وب در یک سایت، نمره‌ی WAB آن وب سایت خواهد بود.

شکل ۱، محاسبه نمره WAB یک وب سایت

را به شکل یک فرمول، بطور مختصر بیان می‌کند. نمره پایین‌تر، به معنای موانع دسترسی پذیری کمتر برای افراد معلول می‌باشد، در حالی که نمره بالاتر، موانع بیشتری را نشان می‌دهد.

نمره صفر، به این معنی است که وب سایت، هیچ راهنمای از دسترسی پذیری وب را نقض نکرده است و نباید هیچ مانع دسترسی برای افراد معلول داشته باشد. از نظر تئوری، فرمول WAB می‌تواند برای محاسبه نمرات WAB بر اساس تمامی ۹۱ نقطه بازرسی در همه اولویت‌های WCAG به کار رود.

$$WAB = \frac{\sum_{i=1}^n \sum_{j=1}^m \left( \frac{w_{ij}}{N_{ij}} \right) (w_{ij})}{N_p}$$

WAB امتیاز  
 کل صفحات وب سایت  
 کل خطاهای یک صفحه وب  
 تعداد خطاها  
 تعداد خطاهای بالقوه  
 وزن خطاها با نسبت معکوس  
 سطح اولویت در WCAG  
 تعداد کل صفحات بررسی شده

### تصویر ۱ - فرمول WAB

اما از آنجا که ما تنها بر آن دسته از نقاط بازرسی که می‌توانند با استفاده از یک روش خودکار ارزیابی شوند توجه داریم، تنها از ۲۵ قاعده، به تفکیک زیر استفاده کردیم: ۵ نقطه بازرسی در اولویت ۱، ۱۳ نقطه بازرسی در اولویت ۲ و ۷ نقطه بازرسی در اولویت ۳. برای شرح جزئیات ۲۵ نقطه بازرسی، به پیوست نگاه کنید.

ما برنامه‌ای به اسم کلونین<sup>۱۰</sup> ساختم که فرمول معیار دسترسی وب را اجرا می‌کند. کلونین، برنامه‌ای بر مبنای جاوا است که از دو قسمت اصلی تهیه شده است: یک خزنده وب<sup>۱۱</sup> و یک ارزیابگر دسترسی<sup>۱۲</sup>. خزنده وب، یک خزنده خودکار سبک وزن می‌باشد که لینک‌ها را برای دیدن صفحات وب دنبال می‌نماید. ما از دیگر خزنده‌های موجود وب استفاده نکردیم، چرا که بیشتر بسیار پیچیده‌تر از آن هستند که بتوان آن‌ها را با کار ساده‌ی ما منطبق نمود. این خزنده می‌تواند به صفحات وب در وب سایت‌های دور دست، دست یابد و تعداد خطاهای بالقوه نقاط بازرسی دسترسی وب را تعیین کند. ارزیابگر دسترسی، خطاهای بالقوه را در مورد ۲۵ نقطه بازرسی WCAG بررسی کرده و نمره دسترسی را محاسبه می‌نماید.

### مثال هایی از ارزیابی دسترس پذیری

در این بخش، به شرح کاربرد معیار ارزشیابی وب سایت هایی که دارای تراکم های مختلف هستند می پردازیم: مقایسه ای از دسترس پذیری وب سایت های مختلف و مقایسه ای از دسترس پذیری وب سایت ها در زمان های مختلف. تحقیق منتشر شده قبلی ما (Zeng & Parmanto, 2004) مثالی در مورد استفاده از روش اندازه گیری WAB به منظور مقایسه طبقات مختلف وب سایت های حاوی اطلاعات بهداشتی برای مصرف کنندگان (آموزش و پرورش، دولت، تجاری، درگاه و جامعه) عرضه می کند.

### دسترس پذیری مجله علمی

مورد اول، یک مطالعه دسترس پذیری وب سایت های مجلات پزشکی سطح بالا می باشد. ما از رتبه بندی ۲۰۰۱ بر اساس عامل تأثیری که توسط نمایه استنادی ISI<sup>۱۳</sup> (نمایه استنادی علوم ۲۰۰۲) گزارش شده است، استفاده نمودیم. مجلاتی را با عامل تأثیر بالاتر از ۱ درصد انتخاب شد. نتیجه اینکه، ۳۸ مجله از مجموع ۱۱۲ مجله، توسط نمایه استنادی علوم (۲۰۰۱) برای پزشکی عمومی رتبه بندی شده اند. تحقیقات پیشین ارزیابی دسترس پذیری وب، تنها صفحه اصلی وب سایت ها (صفحه خانگی) را ارزیابی کرده اند. در حالی که صفحات خانگی معمولاً نماینده خوبی برای دسترس پذیری کامل وب سایت ها نمی باشند. وب سایت مجلات، مثال خوبی از این مورد می باشند.

به جای استفاده محض از صفحه خانگی، ما صفحه خانگی و تمامی صفحات موجود در سطح دوم را ارزیابی نمودیم. تعداد صفحات ارزیابی شده هر وب سایت، بین ۸ تا ۸۹۳ بود. وب سایت هایی را که تنها امکان دسترسی به صفحه اصلی را می دادند (سه وب سایت)، و مجلاتی را که وب سایتی نداشتند (شش مجله) حذف نمودیم، نتیجه، ۲۹ وب سایت مجله برای این مطالعه بود. شکل ۲ امتیاز محاسبه شده برای هر مجله را نشان می دهد. بهترین امتیاز، ۳/۹۸ است که توسط مجله پزشکی آمریکا<sup>۱۴</sup> بدست آمد، درحالی که بدترین امتیاز، ۱۲/۵۵ برای مجموعه مقالات پزشکان آمریکا<sup>۱۵</sup> بدست آمد.

با استفاده از روش اندازه گیری دو بخشی، همه این وب سایت ها، حتی زمانی که از آخرین تابع دقیق (A) استفاده می شود، غیر قابل دسترس به حساب خواهند آمد.

### مطالعه دسترس پذیری وب در طول زمان

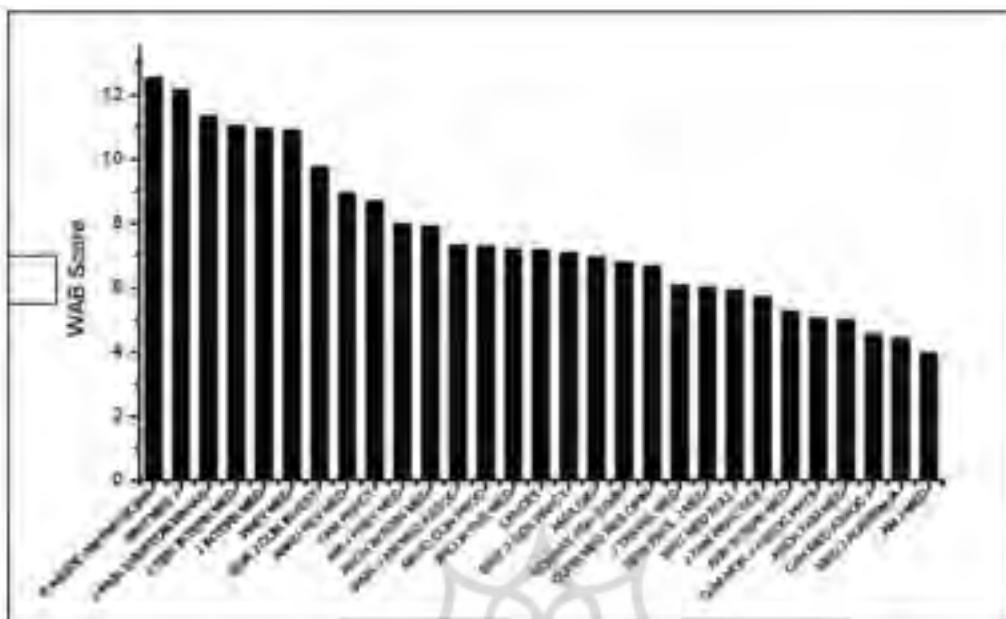
مثال دوم نشان می دهد که چگونه این معیار می تواند برای اجرای یک تحقیق طولی از دسترس پذیری وب، در طول زمان مورد استفاده قرار گیرد. در این مثال، ما دسترس پذیری یک وب سایت را به منظور مشاهده چگونگی تغییر دسترس پذیری آن در طول زمان ارزیابی نمودیم. وب سایت مدیریت غذا و دارو<sup>۱۶</sup> (FDA) را که وب سایت یک آژانس فدرال امریکا است ارزیابی نمودیم (<http://www.fda.gov>). به منظور اجرای این تحقیق، از Wayback Machine، که یکی از خدمات آرشیو اینترنتی است و اینترنت الکسا<sup>۱۷</sup> استفاده نمودیم. آرشیو اینترنتی، در سال ۱۹۹۶ آرشیو کردن وب را که به سرعت در حال تغییر بود آغاز کرد. در تلاشی برای حفاظت از این مواد نمود. سال ۲۰۰۱، که ماشین Wayback برای عموم قابل استفاده شد که به مردم امکان دست یابی و استفاده از ویرایش های مختلف وب سایت های ذخیره شده را داده است، بیش از ۱۰۰ تریلیون بایت اطلاعات با سرعت رشد ۱۲۱۲ تریلیون بایت در هر ماه (Yaukey, 2003) از قبل در آنجا وجود داشت. این کتابخانه اینترنتی از این نوع، به ما امکان می داد که به عقب برگشته و آنچه را که در طول زمان در رابطه با طراحی صفحات وب و دسترس پذیری رخ داده است تجزیه و تحلیل نماییم.

برای هر سال آرشیو شده، نمونه ای از یک وب سایت کامل را که او این نمونه آرشیو شده برای همان سال بود در یک زمان مشخص انتخاب نمودیم. اگر قادر به استفاده از اولین مورد نمی شدیم (به سبب ارتباط ناقص، خطای مسیر نمایه، اخطار فایل یافته نشد و برخی اخطارهای دیگر ماشین Wayback)، مورد آرشیو شده بعدی را برای آن سال به کار می بردیم و به همین ترتیب جلو می رفتیم. در نتیجه، هشت مورد آرشیو شده مجزا را ارزیابی نمودیم. این موارد وب سایت هایی را از سال ۱۹۹۷ تا ۲۰۰۴ نشان می دهند.

نمودار رسم شده در شکل ۳ نشان می دهد که دسترس پذیری برآورد شده وب سایت FDA از سال ۱۹۹۸ تا ۲۰۰۰ و از سال ۲۰۰۱ تا ۲۰۰۳، همانطور که روند افزایش امتیازات مانع دسترس پذیری وب نشان می دهد، بدتر شده است.

سطح امتیاز، از سال ۱۹۹۷ تا ۱۹۹۸ و از ۲۰۰۰ تا ۲۰۰۱ ثابت مانده و یا بهبود کمی دارد. این نمودار همچنین از سال ۲۰۰۳ تا ۲۰۰۴ بهبود قابل ملاحظه ای را در امتیازهای موانع دسترسی نشان می دهد.





شکل ۲. امتیاز های WAB برای ۲۹ مجله سطح بالای پزشکی (امتیاز کمتر، بهتر است).  
 عناوین اختصاری مجلات



شکل ۳. امتیاز WAB برای وب سایت مدیریت غذا و دارو (http://www.fda.gov) : FDA

است استفاده می کند، درگیر قضاوت های ذهنی یا متغیر احتمالات نمی شود. نتایج اندازه گیری به طور عینی تا حدودی وضعیت دسترس پذیری محتوای صفحه وب را منعکس می کند. شیوه های سستی

سنجش اعتبار معیار پایایی معیار از آنجا که این روش اندازه گیری، از داده هایی که مستقیماً توسط فرآیند ماشینی خودکار بدست آمده

پایایی (۵ درجه بندی داخلی، آزمون- باز آزمون، فرم های موازی، همبستگی درونی) برای معیار ما کاربردی ندارند.

### اعتبار معیار

استاندارد طلایی: برای ارزیابی معیار عددی، یک استاندارد طلایی باید به کار گرفته شود. منظور از یک استاندارد طلایی، استاندارد مرجعی برای ارزیابی یک معیار تشخیص بدیع است که در این تحقیق به امتیاز مانع دسترس پذیری وب اطلاق می گردد.

چندین روش اندازه گیری پیشنهادی، می توانند به عنوان استاندارد طلایی در این تحقیق به کار گرفته شوند. اولاً، ما می توانیم از افراد معلول برای قضاوت درباره دسترس پذیری یک صفحه وب استفاده نماییم. اگرچه این شیوه ی ایده آلی است، اما این رویکرد به نظر غیر عملی می رسد. خود افراد معلول، با در نظر گرفتن انواع و سطوح مختلف ناتوانی ها گروه بسیار متنوعی را تشکیل می دهند. شرایط لازم برای دسترس پذیری از هر زیر گروه، بسیار ویژه و اغلب متعارض می باشند. مثال بارز آن می تواند یک صفحه وب متن محور باشد که برای افرادی با نقص بینایی بسیار قابل دسترسی است در حالی که برای افرادی با ناتوانی یادگیری غیر قابل دسترسی است (Bohman, 2003).

روش دیگر، استفاده از ارزیابی جامع صفحات وب به طور دستی با پیروی از استاندارد های دسترس پذیری وب می باشد. WAI الگویی را برای ارزیابی جامع سطح دسترس پذیری یک صفحه وب را منتشر ساخت (WAI, 2002). این الگو شامل مراحل گوناگون، ابزارهای متنوع و کنترل های دستی فراوان می باشد. چنین روش اندازه گیری بسیار بالا است و برای تعداد زیادی از صفحات، به طور باز دارنده ای گران می باشد. گزینه سوم در روش سنجش استاندارد طلایی، این است که بر بعضی از موارد معتبر تکیه شود.

از آنجا که راهنماهای دسترس پذیری محتوای وب به منظور خدمات رسانی به طیف وسیعی از معلولین طراحی شده است، نامزد خوبی برای استاندارد طلایی دسترس پذیری به حساب می آید. همانطور که در بخش پیش مطرح شد، چارچوب های WAI قالب های تطابق WCAG، از سطح A تا سطح AAA به همان شکل که در بخش پیشین مطرح شد را به منظور بالا بردن دسترس پذیری وب معرفی نموده

است. تهیه کنندگان محتوا می توانند به منظور اثبات ادعای تطابق با سطوح ویژه ی WCAG 1.0 از این چارچوب ها در سایت های خود استفاده نمایند. پس از آن که تهیه کنندگان محتوا صفحات وب خود را با نقاط بازرسی راهنماهای دسترس پذیری محتوای وب منطبق ساختند، می توانند چارچوب WAI را به صفحات خود بیفزایند.

سطح تطابق، نوع چارچوبی را که می تواند بکار ببرند تعیین می کند. از آنجا که چارچوب های WAI خودشان دارای تصاویری هستند که در صفحه HTML جاسازی شده اند، معمولاً دارای متن جایگزینی هستند که به آنها متصل شده است. به عنوان مثال، متن جایگزین برای چارچوب انطباق با سطح A، عبارت است از نشانه تطابق سطح A، راهنمای محتوای وب و دسترس پذیری W3C-WAI ویرایش 1.0. به عنوان پیش فرض، آیکن تطابق به یک صفحه ی واحدی ارجاع می دهد، مگر اینکه یاد داشت دامنه بطور آشکارا مدعی چیز دیگری باشد.

نظام قالبی WAI، کاندیدای بالقوه ای برای یک روش سنجش استاندارد طلایی است. این سیستم جامع است و حیطة وسیعی از معلولیت ها را در بر می گیرد و از نظر هزینه، کار آ است. با این همه سیستم چندین اشکال دارد که ممکن است نتایج تحقیق را تحت تأثیر قرار دهد. سیستم قالبی، یک سیستم خود امتیاز بندی<sup>۱۸</sup> است. همانطور که در قسمت پیشینه بحث شد، این سیستم خود - امتیاز بندی کامل نیست. با وجود اینکه که سیستم قالبی اشکالاتی دارد، باز هم یک استاندارد طلایی مطلوب برای مطالعه روش سنجش می باشد. هر چند که سیستم قالبی کامل نیست، اما نشان می دهد که تهیه کننده ی محتوا برای رفع نمودن موانع دسترس پذیری از وب سایت کار قابل توجهی انجام داده است. بعلاوه، عمومیت وب باعث می شود صفحات وبی که از این قالب ها استفاده می کنند، دائماً توسط عموم مورد مذاقه قرار گیرند و در نتیجه، با هر ناهمخوانی با نقاط بررسی WAI، اظهار نارضایتی و شکایتی علیه آن ها صورت خواهد گیرد. ما از موتور جستجوی گوگل برای بدست آوردن صفحات وبی که به عنوان یک استاندارد طلایی عمل می کنند، استفاده نمودیم. این صفحاتی که توسط موتور جستجو یافت شدند به منظور تأیید وجود و نوع قالب دسترس پذیری وبی که آنها مورد استفاده قرار دادند آزمایش می شوند. ما صفحات خانگی ۵۰۰ وب سایت رتبه بالا را که توسط یک موتور جستجوی دیگر - الکسا - تهیه شده بود، به

عنوان یک گروه منفی بدون قالب WAI، به کار بردیم. تک تک صفحات وب موجود در گروه منفی برای تأیید فقدان قالب WAI بیشتر مورد آزمایش قرار گرفتند. نمونه‌ای از وب سایت هایی که در این تحلیل مورد استفاده قرار گرفت همانند آنهایی است که در جدول ۱ آورده شد.

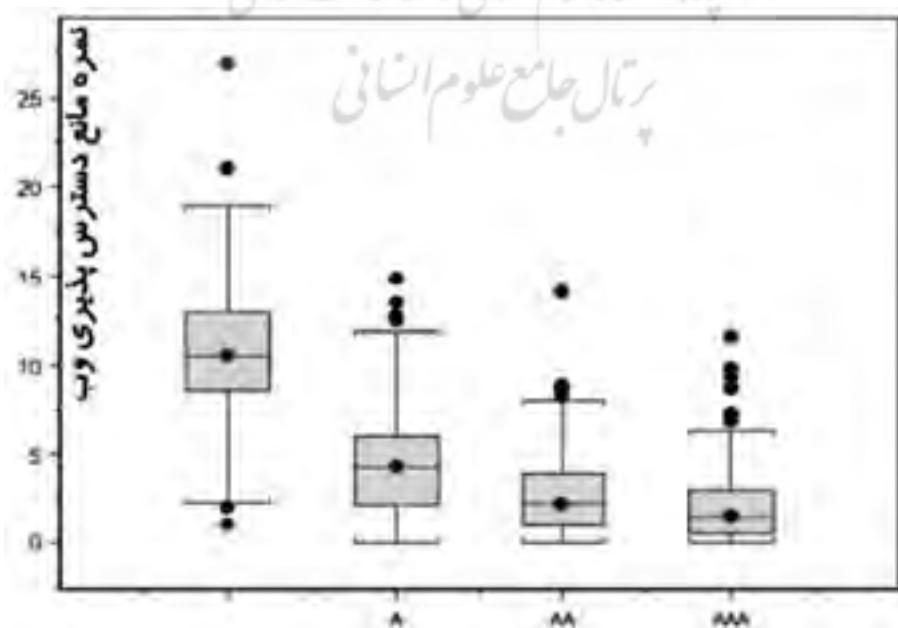
### نتایج

نتایج کاربرد معیار در وب سایت هایی که به عنوان استاندارد طلایی جمع آوری شدند، در شکل ۴ ارائه شده است. نتایج نشان می دهد که معیار WAB همواره نماینده خوبی است از وب سایت هایی که دسترس پذیری آن ها برآورد شده است. به طور متوسط، وب سایت هایی که خود را در سطح AAA در نظر گرفته بودند نسبت به آن هایی که خود را AA در نظر گرفته بودند امتیاز مانع دسترس پذیری وب بهتری داشتند و این ها نیز امتیاز بهتری نسبت به وب سایت های A داشتند که باز به نوبه خود امتیازشان بهتر از وب سایت های امتیاز بندی نشده بود.

معیار WAB، «درجه های» پیوسته ای از دسترس پذیری برآورد شده را عرضه می کند. امتیازهای متوسط وب سایت های A، AA، AAA و وب سایت های امتیاز بندی نشده به ترتیب ۲/۰۲، ۲/۷۴، ۲/۴۷ و ۱۰/۵ می باشند. شکل ۴، همچنین تعدادی از وب سایت ها را نشان می دهد که جزء دسته های امتیاز بندی شده (سطح AAA، AA و A) هستند و امتیاز آن ها از متوسط

امتیاز وب سایت های امتیاز بندی نشده بدتر بود. این وب سایت های بیرونی در بالای Box و نمودار Whisker بصورت نقطه هایی نمایش داده شده است که بیانگر داده های خارج از محدوده اطمینان ۹۵٪ می باشند. همچنین وب سایت های امتیاز بندی نشده ای وجود داشتند که نسبت به متوسط امتیاز وب سایت های AAA بهتر امتیاز گرفتند که توسط نقاط زیر نمودار (وب سایت های) امتیاز بندی نشده، نشان داده شده اند. این شکل نشان می دهد که معیار WAB قادر است وب سایت ها را از یکدیگر بر اساس دسترس پذیری برآورد شده شان در سراسر سطوح مختلف طیف دسترس پذیری تفکیک سازد. ما آزمایش کردیم که چگونه امتیاز وب سایت های سطح A، AA، AAA و وب سایت های امتیاز بندی نشده از یکدیگر جدا هستند. عملکرد معیار سنجش در پیش بینی هر یک از دسته وب سایت ها با استفاده از منحنی مختصات عملیاتی دریافت کننده<sup>۱۹</sup> (اعم از اینکه وب سایت به AA، AAA یا A متعلق باشد) محاسبه خواهد شد (Egan, 1975).

منحنی ROC برای آزمایش اعتبار این معیار بسیار مناسب تر است تا آزمایش ساده میانگین اختلاف نمرات با استفاده از روش تحلیل واریانس یکسویه (Frideman & Wyatt, 1997). منحنی ROC بطور معمول برای ارزیابی توانایی یک پیش بین در تمایز بین دو نتیجه ی ممکن بکار می رود.



سطوح مطابقت WCAG 1.0

شکل ۴. امتیاز های WAB وب سایت ها در میان سطوح مختلف مطابقت.

(امتیاز بندی نشده از A، A از AA و AA از AAA). تمایز بین گروه امتیاز بندی نشده و گروه A قوی ترین است، حال آنکه تمایز بین سطوح AA و AAA ضعیف ترین است.

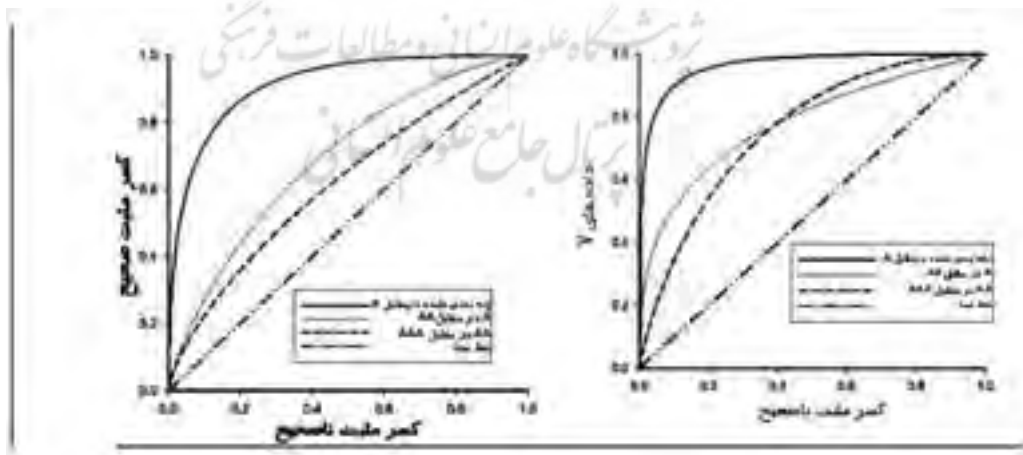
جدول ۲ امتیاز سطح زیر منحنی معیار را در جداسازی گروه های مختلف امتیاز بندی های دسترس پذیری نشان می دهد. نتایج تمایز آشکاری را بین وب سایت های غیر قابل دسترس و وب سایت هایی با امتیاز بندی سطح A نشان می دهد. امتیاز ۹/۱۷ نمره سطح زیر منحنی برابر ۹/۱۷ است و تمایز بین وب سایت های امتیاز بندی نشده با وب سایت سطح A بسیار قابل توجه و معنی دار است (۰/۰۰۰۱ < مقدار P). تمایز بین وب سایت های سطح AAA و وب سایت های امتیاز بندی نشده با نمره ۹/۷ و (۰/۰۰۰۱ < مقدار P) حتی بیشتر است. تمایز ضعیف تر بین AAA و AA نیز قابل توجه است (۰/۰۰۰۱ < مقدار P).

پس از آن برای مقایسه عملکرد طرح وزن دهی ساده ای که در معیار WAB خود به کار بردیم با یک روش پیچیده تر درخت تصمیم گیری که از الگوریتم یادگیری ماشینی C5.0 که یادگیری آن بر اساس مجموعه داده ها است (Quilan, 1993)

رسم یک منحنی ROC نقاطی را که روی یک کسر مثبت صحیح (TFP) قرار دارند و نقاطی را که روی یک کسر مثبت ناصحیح (FPF) قرار دارند و بر نقاط تقاطع در طول معیار قرار دارند، به یکدیگر متصل می کند. سطح زیر منحنی<sup>۲۰</sup> (AUC) منعکس کننده توان تشخیص آزمون است. تفکیک کامل بین دو گروه دارای نمره سطح زیر منحنی ۱ بوده و منحنی در امتداد محور X ها و بالای محور Y ها است. در مقابل، تفکیک ناپذیری کامل دارای نمره سطح زیر منحنی برابر با ۰/۵ بوده و منحنی مربوطه در امتداد خط راست است.

از خصوصیات دیگر منحنی ROC این است که یک نقطه تقاطع معین یا نقطه معیار را می توان روی منحنی یافت. ما نرم افزار ROCKIT<sup>۲۱</sup> که دارای حساسیت و خصوصیت ترجیحی می باشد را از دانشگاه شیکاگو برای تجزیه و تحلیل منحنی ROC بکار بردیم (University of Chicago, 2004).

شکل ۵ منحنی های ROC را که از نقاط برشی متفاوت با امتیاز WAB برای استاندارد طلایی وب سایت ها ترسیم شده اند نشان می دهد. منحنی ها نشان دهنده توان معیار در سنجش میزان جداسازی سطوح همجوار در گروه های امتیاز بندی وب سایت هاست



تصویر ۵. منحنی ROC از نمره مانع دسترس پذیری وب، وب سایت های سطح A را از رتبه بندی نشده، سطح AAA از سطح A و سطح AAA از سطح AAA جدا می سازد.

تصویر ۶. عملکرد ROC از الگوریتم یادگیری ماشینی C5.0، در جداسازی وب سایت های سطح A از رتبه بندی نشده، سطح AA از سطح AAA و سطح AAA از سطح AAA.



را کاهش داده و حداقل تعداد رکورد ها برای هر زیر شاخه روی ۵ تنظیم شد. مجموعه قواعدی را به عنوان برون‌داد مدلی که توسط الگوریتم C5.0 ایجاد شده بود، برگزیدیم. عمق درخت ها برای هر سه مجموعه داده ها، دو سطح است. همچنین از منحنی ROC برای ارزیابی عملکرد درخت تصمیم‌گیری در هر یک از مجموعه داده ها استفاده نمودیم.

شکل ۶ منحنی های ROC را نشان می‌دهد که از نقاط برش مختلف برای نمرات وزن دهی پیچیده ای که توسط الگوریتم یادگیری ماشینی C5.0 برای همان وب سایت ها بدست آمده بودند ترسیم شده است. الگوریتم C5.0 در مقایسه با نمره WAB در جداسازی گروه های نزدیک دسترس پذیری، بویژه در جداسازی AA و AAA به خوبی عمل می‌کند. جدول ۳ نمره سطح زیر منحنی از یادگیری ماشینی C5.0 را در جداسازی سطوح مختلف دسترس پذیری نشان می‌دهد. همانطور که انتظار می‌رفت، عملکرد یادگیری ماشینی، بهتر از طرح وزن دهی ساده‌ی به کاررفته در معیار WAB بود. جدول ۴ مقادیر سطح زیر منحنی را میان امتیاز مانع دسترس پذیری و روش C5.0 مقایسه می‌کند. این جدول نشان می‌دهد که عملکرد معیار WAB در جداسازی دسته های رتبه بندی شده از رتبه بندی نشده ها به خوبی روش پیچیده‌ی C5.0 می‌باشد. تفاوت ها در مقادیر سطح زیر منحنی جداسازی دو دسته از سه دسته (AA از رتبه بندی نشده و AAA از رتبه بندی نشده) چشمگیر نمی‌باشد.

جدول ۲. AUC (مساحت زیر نمودار) امتیاز WAB در جداسازی وب سایت های AA، AAA و A از وب سایت های رتبه بندی نشده.

رتبه بندی نشده	A	AAA	AA-AAA	AAA	رتبه بندی نشده
0/972	0/917	0/899	0/813	0/982	0/972

نکته. به این معناست که تمامی سطح زیر نمودار به طور چشمگیری با ۰/۵ متفاوت هستند.

جدول ۳. مقدار AUC (مساحت زیر منحنی) روش یادگیری ماشینی (C5.0) در جداسازی AA، AAA و A از وب سایت های رتبه بندی نشده.

رتبه بندی نشده	A	AAA	AA-AAA	AAA	رتبه بندی نشده
0/983	0/962	0/797	0/799	0/983	0/983

نکته. برای همه سطح زیر نمودار، مقدار  $p > 0.01$ .

از روش های یادگیری ماشینی استفاده کردیم. تفاوت اصلی بین نمره WAB و روش فراگیری ماشینی این است که روش نمره WAB از امتیاز رتبه بندی ساده‌ی از پیش تعیین شده‌ای استفاده می‌کند که با سطح اولویت خطاها نسبت عکس دارد (نمره ۳ برای اولویت ۱ نمره ۲ برای اولویت ۲ و نمره ۱ برای اولویت ۳).

یادگیری ماشینی از مجموعه داده ها اطلاع می‌گیرد و مقادیر مطلوب را به هر یک از ۲۵ نقطه بازرسی اختصاص می‌دهد. هدف از این مقایسه پی بردن به میزان درستی مقایسه میان طرح وزن دهی ساده و طرح وزن دهی پیچیده مطلوب می‌باشد. ما برای ساختن مدل C5.0 از نرم افزار کلماتین متعلق به شرکت SPSS استفاده کردیم. نسبت خطاها (خطاهای واقعی به خطاهای بالقوه) برای هر ۲۵ نقطه بازرسی خودکار، به عنوان مدل به کار گرفته شد. متغیر برون‌داد مدل، سطح تطابق WCAG است: رتبه بندی نشده، AAA، AA، A. برخلاف نمره WAB، الگوریتم درختی تصمیم‌گیری، فقط برای جداسازی دو سطح تطابق در یک زمان مورد استفاده قرار گرفت (بین رتبه بندی نشده (ها) A و A- AA و غیره). ما یک روش اعتباریابی و سه برابر را برای ارزیابی دقت مدل درخت تصمیم‌گیری به کار بردیم: تقریباً دو سوم وب سایت ها را برای ساختن مدل و یک سوم را برای آزمایش مدل به کار بردیم. در ضمن اعتباریابی، نسبت وب سایت ها در هر سطح تطابق در دسته بندی مجموعه داده ها در پارامترهای ساخت درخت تصمیم‌گیری شدت ۰/۷۵.

و قابلیت اطمینان طرح وزن دهی ساده آن را جالب تر می سازد.

### محدودیت های معیار

معیار دسترس پذیری که ما تهیه نمودیم، برای روش های سنجش عینی و سیستماتیک از دسترس پذیری وب در نظر گرفته شده است. این معیار برای بررسی کنترل دسترس پذیری یک وب سایت خاص با هدف اصلاح دسترس پذیری یا بهبود آن مناسب نیست. این معیار همچنین به عنوان مقیاس نماینده دسترس پذیری در نظر گرفته شده است، نه مقیاسی واقعی که نیازمند بررسی دستی و تشخیص انسانی است. این معیار محل مانع را که می تواند بر قابلیت استفاده از وب سایت تاثیر گذارد، مورد توجه قرار نمی دهد. (هرقدر جایگاه مانع وب در سلسله مراتب وب سایت بالاتر باشد، مشکلات بالقوه در استفاده از آن بیشتر است). محل مانع می تواند در بازنگری های آینده ی معیار مورد توجه قرار گیرد.

### نتیجه گیری

ما معیار جدیدی را برای سنجش دسترس پذیری وب پیشنهاد نمودیم که نیازمندی های یک روش سنجش برای تحقیق علمی را برآورده می سازد. این معیار می تواند برای ارزیابی عینی و مقایسه دسترس پذیری بین وب سایت های مختلف، گروه های مختلف وب سایت و وب سایت های مختلف یا گروه های وب سایت در زمان های مختلف به کار رود.

این معیار ساده، به خوبی با روش پیچیده یادگیری ماشینی قابل مقایسه است. بر این باوریم که وجود یک معیار عینی، درهای رویکرد علمی را به مطالعات دسترس پذیری وب خواهد گشود. این مطالعه، با نشان دادن امکان پذیری ارزشیابی خودکار دسترس پذیری وب، مضامین مهمی برای معیارهای سنجش خودکار مشابه، در بر دارد.

### قدر دانی

بخشی از این تحقیق توسط کمک پژوهشی شماره 42-60-IO2013 مدیریت مخابرات و اطلاعات ملی و نیز کمک پژوهشی شماره H133AO21916 موسسه ملی تحقیقات معلولین و باز توانی مورد حمایت قرار گرفت. نویسندگان مایلند از سجاریف احمد به خاطر تهیه مقیاس کلونین که در این تحقیق مورد استفاده قرار گرفت تشکر نمایند و نیز از استفانی هکت به خاطر انجام تجزیه و تحلیل ها در مطالعه طولی قدردانی می نمایند.

اگرچه تفاوت در مقادیر سطح زیر منحنی برای جداسازی A از رتبه بندی نشده ها، قابل توجه است، اما عملکرد معیار WAB نیز عالی است (۰,۹۱۷). این مفهوم با کاربرد روش تحلیل آماری دو متغیره با این فرض باطل که دو منحنی ROC همانند می باشند، محاسبه شده است. با استفاده از روش تحلیل آماری دو متغیره سطح معنی داری محاسبه شد. در این محاسبات فرض صفر این بود که منحنی های ROC یکسان هستند. این تحلیل با استفاده از نرم افزار ROCKIT انجام می شود. (Hanley و McNeil) پیشینه آماری مفصلی آماری را در محاسبه سطح زیر منحنی و آزمایش معنی داری در مقایسه دو منحنی ROC عرضه نمودند.

الگوریتم C5.0 به میزان قابل توجهی از امتیاز WAB در جداسازی AAA و AA بهتر است. این الگوریتم نیز در جداسازی A و AA نیز بهتر است، اگرچه عملکرد WAB نیز خوب است (۰,۶۹). ما اجزای درخت تصمیم گیری را که با الگوریتم C5.0 ساخته شد، به منظور آزمایش چگونگی تفاوت آن از WAB در ارزیابی مقادیر نقاط بازرسی دسترس پذیری از هم جدا نمودیم.

امتیاز WAB همه ی ۲۵ نقطه بازرسی را در نظر می گیرد، در صورتیکه الگوریتم C5.0 تنها زیر مجموعه ای از ۲۵ نقطه بازرسی را انتخاب می کند. تعداد نقاط بازرسی در امتیاز WAB برای هماهنگی با اولویت های ۲ و ۳ به ترتیب ۱۳ و ۵ (نسبت ۱۳ به ۵ است) می باشد. تعداد نقاط بازرسی انتخاب شده توسط الگوریتم C5.0 برای هماهنگی با اولویت های ۲ و ۳، ۴ و ۳ (نسبت ۴ به ۳ است) می باشد.

این امر ممکن است توضیح این نکته باشد که چرا الگوریتم درخت تصمیم گیری با دادن وزن بیشتر به نقطه بازرسی ها در اولویت ۳ (نقاط بازرسی که AA و AAA را جدا می سازد) بطور قابل توجهی در جداسازی AA و AAA بهتر عمل می کند. توزیع نقاط بازرسی ایجاد شده توسط C5.0 در جدول ۵ نشان داده شده اند.

نتایج نشان می دهند که روش وزن دهی ساده که در نمره WAB به کار رفته است در مقایسه با روش پیچیده تر درخت تصمیم گیری در کارهای جداسازی مهم، خوب عمل می کند (جداسازی گروه های رتبه بندی شده از رتبه بندی نشده ها). از آن جایی که درخت تصمیم گیری پیچیده تر است و طرح های وزن دهی مختلفی را در رابطه با مجموعه داده های مختلف ارائه می کند، سادگی

- 1-Web Content Accessibility Guidelines
- 2- World Wide Web Consortium
- 3- Sullivan & Matson
- 4-Trace Research and Development Center
- 5-Web Accessibility Initiative
- 6-Bobby approved
- 7-prohibitively expensive
- 8- McMullin
- 9-Web Accessibility Barrier
- 10-Kelvin
- 11-Web crawler
- 12-accessibility evaluator
- 13-ISI Citation Index
- 14-American Journal of Medicine
- 15- Proceedings of the Association of American Physicians
- 16-Food and Drug Administration
- 17- Internet Archive and Alexa Internet
- 18-self-rated
- 19-Receiver Operating Characteristics (ROC) Curve
- 20- Area Under the Curve

۲۱ - یک نرم افزار آماری برای ترسیم و تجزیه تحلیل منحنی ROC که توسط دامشگاه شیگاه نوشته شده است (مترجم).

پژوهشگاه علوم انسانی و مطالعات فرهنگی

## منابع و ماخذ

- Australian Human Rights & Equal Opportunity Commission. (1997). Disability standards and guidelines. Retrieved December 2, 2003, from [http://www.hreoc.gov.au/disability\\_rights/standards/standards.html](http://www.hreoc.gov.au/disability_rights/standards/standards.html)
- Bergman, M.K. (2001). The deep web: Surfacing hidden value. *The Journal of Electronic Publishing*, 7(1). Retrieved from <http://www.press.umich.edu/jep/07-01/bergman.html>
- Bohman, P.R. (2003). Visual vs. cognitive disabilities. Retrieved December 8, 2003, from [http://www.webaim.org/techniques/articles/vis\\_vs\\_cog](http://www.webaim.org/techniques/articles/vis_vs_cog)
- Deaukantas, P. (2002). Think tank report: Federal web sites need better accessibility. Retrieved March 4, 2004, from [http://www.gcn.com/vol1\\_no1/s508/19757-1.html](http://www.gcn.com/vol1_no1/s508/19757-1.html)
- Deaukantas, P. (2002). Think tank report: Federal web sites need better accessibility. Retrieved March 4, 2004, from [http://www.gcn.com/vol1\\_no1/s508/19757-1.html](http://www.gcn.com/vol1_no1/s508/19757-1.html)
- Dellavalle, R.P., Hester, E.J., Heilig, L.F., Drake, A.L., Kuntzman, J.W., Graber, M., et al. (2003). INFORMATION SCIENCE: Going, going, gone: Lost internet references. *Science*, 302(5646), 787-788.
- Dhyani, D., Wee Keong, N., & Bhowmick, S.S. (2002). A survey of web metrics. *ACM Computing Surveys*, 34(4), 469-503.
- Egan, J.P. (1975). *Signal detection theory and ROC analysis*. New York: Academic Press.
- Emery, G.R. (2002). Survey: Agency web sites make progress, still have for to go. Re-

trieved March 4, 2004, from [http://www.washingtontechnology.com/news/1\\_1/daily\\_news/18859-1.html](http://www.washingtontechnology.com/news/1_1/daily_news/18859-1.html)

Friedman, C.P., & Wyatt, J.C. (1997). *Evaluation methods in medical informatics*. New York: Springer.

Hanley, J.A., & McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29-36.

Hanley, J.A., & McNeil, B.J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*, 148(3), 839-843.

Ivory, M.Y., & Hearst, M.A. (2001). The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys*, 33(4), 470-516.

Ivory, M.Y., & Hearst, M.A. (2002). Improving web site design. *IEEE Internet Computing*, 6(2), 56-63.

Jackson-Sanborn, E., Odess-Harnish, K., & Warren, N. (2002). Web site accessibility: A study of six genres. *Library Hi Tech*, 308-317.

McMullin, B. (2004, July). Users with disability need not apply? Web accessibility in Ireland. *First Monday*, 9(7). Retrieved from [http://www.FirstMonday.org/issue9\\_7/marincu/](http://www.FirstMonday.org/issue9_7/marincu/)

Murray, B.H., & Moore, A. (2000). Sizing the Internet. Retrieved December 2, 2003, from [http://www.cyveillance.com/web/downloads/Sizing\\_the\\_Internet.PDF](http://www.cyveillance.com/web/downloads/Sizing_the_Internet.PDF)

Paciello, M.G. (2000). *Web accessibility for people with disabilities*. Berkeley, CA: CMP Books.

Quinlan, J.R. (1993). *C4.5: Programs for machine learning*. San Mateo, CA: Morgan Kaufmann Publishers. Research Initiative for Network and Communication

Engineering (RINCE) and Association BrailleNet. (2003). Des sites web français de la vie quotidienne sont inaccessibles aux personnes handicapées [Popular French Web sites are inaccessible to people with disabilities]. Retrieved December 1, 2003, from [http://brailletnet.org/eval\\_30sites\\_France\\_10sept2003.htm](http://brailletnet.org/eval_30sites_France_10sept2003.htm)

Science Citation Index. (2002). Philadelphia: Thomson ISI.

SPSS Inc. (2003). *Clementine 7.0*. Retrieved September 1, 2003, from <http://www.spss.com>

Stower, G. (2002). The state of federal websites: The pursuit of excellence. Arlington, VA: The PricewaterhouseCoopers Endowment for the Business of Government.

Sullivan, T., & Matson, R. (2000, November). Barriers to use: Usability and content accessibility on the web's most popular sites. Paper presented at the Proceedings of CUU 2000, ACM Conference on Universal Usability, Arlington, VA.

Tufte, E.R. (1997). *Visual explanations: Images and quantities, evidence and narrative*. Cheshire, CT: Graphics Press.

University of Chicago. (2004). *ROCKIT* [Computer program]. Retrieved January 10, 2004, from <ftp://random.bsd.uchicago.edu/roc/ibmpc/>

Watch.re Corp. (2004). Bobby, Watch.re Corporation. Retrieved June 10, 2004, from <http://bobby.watch.re.com/>

Web Accessibility Initiative (WAI). (1999). *Web Content Accessibility Guidelines 1.0*. Retrieved August 1, 2003, from <http://www.w3.org/TR/WCAG10/>

Web Accessibility Initiative (WAI). (2002). *Evaluating web sites for accessibility*. Retrieved December 8, 2003, from <http://www.w3.org/WAI/eval/Overview.html>

Weiss, R. (2003, November 24). On the web, research work proves ephemeral electronic archivists are playing catch-up in trying to keep documents from landing in history's dustbin. *Washington Post*, p. A08.

West, D.M. (2001). *WMRC global E-government survey*. Retrieved December 2, 2003, from <http://www.insidepolitics.org/egovt01int.html>

Yaukey, J. (2003). Archive site preserves earliest web pages, from <http://www.gannett-online.com/e/trends/15000566.html>

Zeng, X., & Parmanto, B. (2004). Web content accessibility of consumer health information web sites for people with disabilities: A cross sectional evaluation. *Journal of Medical Internet Research*, 6(2), e19. Retrieved from <http://www.jmir.org/2004/2/e19/>



پیوست

نقاط بازرسی به کار رفته در تگرات WAB

اولویت WAI	ID	نقطه بازرسی	تعیین شماره خطاهای بالقوه
1	g 9	برای تمام تصاویر متن معادل تهیه کنید	نام عناصر <img>
1	g 21	برای هر یک از اینلین ها متن معادلی تهیه کنید	نام عناصر <applet>
1	g 20	برای هر شیئی متوالی معادلی تهیه کنید	نام عناصر <object>
1	g 10	برای تمام کلیدهای تصویری موجود در فرم ها متن معادل تهیه کنید	نام عناصر <input type="image">
1	g 340	برای تمام نقاط مهم نقشه ها (نواحی) متن معادل تهیه کنید	نام عناصر <area>
1	g 38	هر قابی نباید به یک پرونده ای سی ام ال ارجاع داده شود	نام عناصر <frame>
1	g 39	به هر قاب عنوانی بدهید	نام عناصر <frame>
2	g 271	این عبارت DOCTYPE از یک شناسگر متنی عمومی استفاده کنید	1*
2	g 104	جای اندازه و جای نظمی (بیشتر) از اندازه و جای نظمی (مقادیر) استفاده کنید	نام عناصر <table>، <th>، <td> و <frame>
2	g 2	جای سرخطها را با یک نامی مشخص کنید	نام عناصر سرخطها
2	g 37	هنگام به کار بردن قاب یک متن بدون قاب هم تهیه کنید	نام عناصر <frameset>
2	g 4	از به کار بردن متون چشمک زن که با عنصر BLINK درست می شوند پرهیز کنید. از به کار بردن متون سیال که با عنصر MARQUEE ساخته می شوند پرهیز کنید	مانند تعداد خطاهای واقعی <sup>۱</sup>
2	g 5	کاری نکنید که صفحه ای خود خود تراز شود	مانند شماره خطاهای واقعی <sup>۱</sup>
2	g 33	کاری نکنید که صفحه ای به یک url جدید هدایت شود	1*
2	g 254	مطمئن شوید که کنترل کلیدها و سایر اقدامات با ماوس کار می کنند	شماره اداره کننده <sup>۲</sup> و تعداد برای صفحه <sup>۳</sup> کلید و ماوس
2	g 269	کنترل کننده های فرم ها و سرچسب های آن ها را با یک عنصر LABEL مرتبط سازید	شماره عنصر شکلی مانند <input>، <select> و <textarea>
2	g 34	عبارت پیوندی را شوری بسازید که هنگامی که مفهوم باشد	شماره عناصر <a>
2	g 265	هنگامی که پیوندها به urlهای مختلف اشاره دارند، بوی از یکبار از یک عبارت پیوندی استفاده نکنید	شماره عناصر <a>
2	g 273	عنوان یک مدرک را نیز بگنجانید	1*
2	g 273	لبته تصاویر مربوط به client شامل پیوندی است که در هیچ جای دیگر صفحه ارائه نشده	شماره عناصر <area>
3	g 125	زبان متن را شناسایی کنید	1*
3	g 31	برای جداول خلاصه ای تهیه کنید	شماره عناصر <table>
3	g 109	نویسه های پیش فرض place-holding را در دستنظیمی های ویرایش و نواحی مختلفی به کار نبرید	شماره عناصر <input type="text">، <select> و <text area>
3	g 35	پیوندهای هماور را با چیزی بیش از فاصله های سفید از یکدیگر جدا کنید	شماره پیوندها

۱ این عنصر فقط یکبار در یک صفحه قابل دسترسی ظاهر می شود.  
 ۲ این عنصر هرگاه در صفحه وبی ظاهر شود، به عنوان خطای راهنمای دسترسی پذیر می شود.  
 ۳ این عنصر هرگاه در صفحه وبی ظاهر شود، به عنوان خطای راهنمای دسترسی پذیر می شود.