Scoring the Accessibility of Websites

Jared Smith
(WebAIM, Associate Director)

Christoph Rump
(Accenture, Test Engineering Manager)

Agenda

1. Introduction

Our Path to a Scoring Methodology

The methodology was collaboratively created by WebAIM at Utah State University (USU) and different business units at Accenture (ACN) with the goal to assess the accessibility of websites using the "Web Content Accessibility Guidelines" (WCAG).

Automated accessibility data is often insufficient to effect change in our respective clients

In-depth manual testing can be difficult and expensive

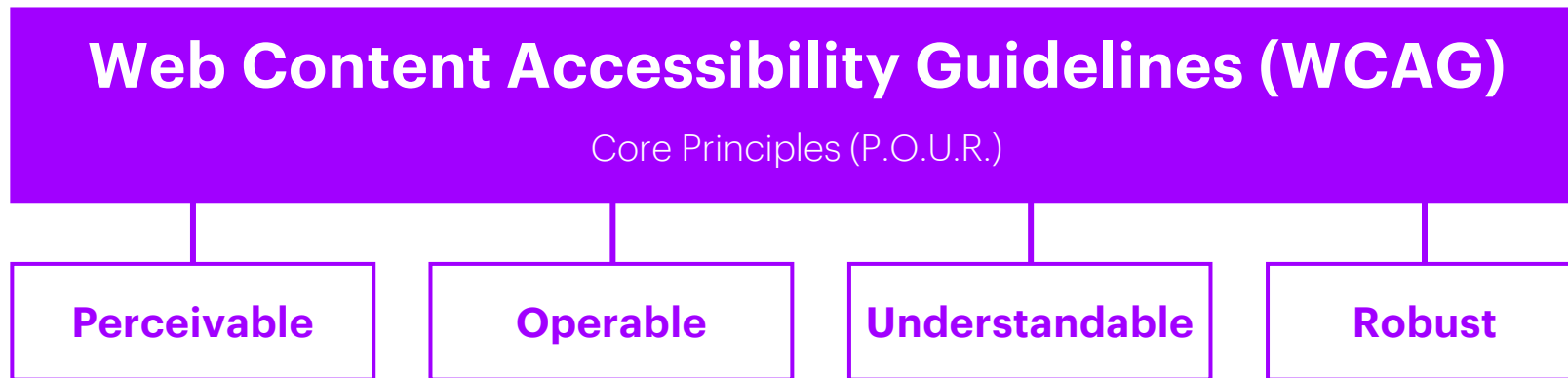Accessibility test data is often descriptive, but not overly prescriptive. Where do we start?

WCAG conformance testing does not always measure human impact

Could we create a methodology to provide automated data, manual testing, and human impact?

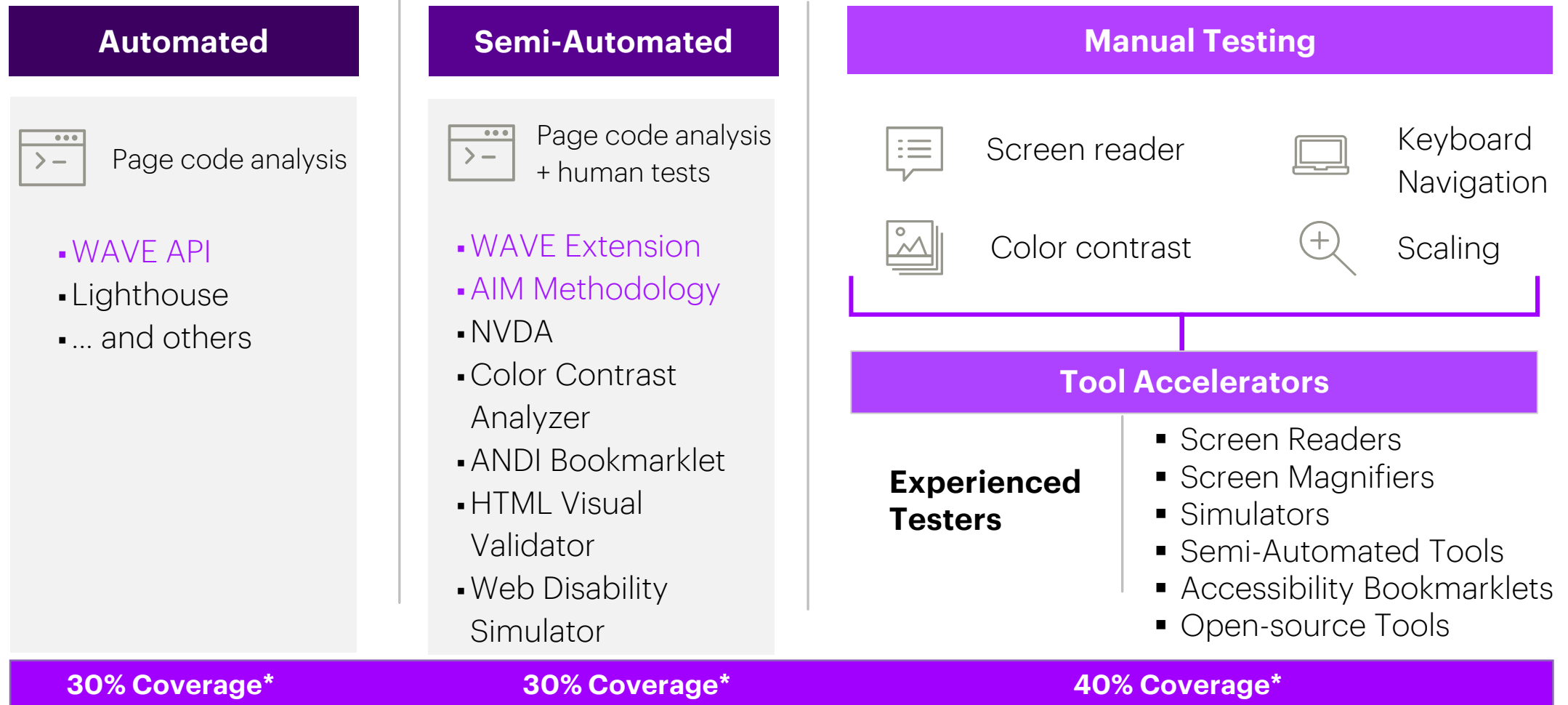Creation of the AIM methodology with normalized scoring, by using the WebAIM Million

WebAIM
web accessibility in mind

accenture

2. WCAG Automation Coverage

Standards and guidelines provide measures for documenting accessibility

EU Mandate	USA: Legal & Financial Risk	Country-specific Guidelines
Industry-specific regulations (e.g. LS)	Corporate Social Responsibility	ISO Norm

Web Content Accessibility Guidelines (WCAG)

Core Principles (P.O.U.R.)

Perceivable	**Operable**	**Understandable**	**Robust**

78 total criteria organized as **13** guidelines under **4** principles.

Assistive Technologies

WebAIM
web accessibility in mind

accenture

Accessibility testing is the practice of measuring web and mobile app usability for users with disabilities

Automated

Page code analysis

- WAVE API
- Lighthouse
- … and others

Semi-Automated

Page code analysis + human tests

- WAVE Extension
- AIM Methodology
- NVDA
- Color Contrast Analyzer
- ANDI Bookmarklet
- HTML Visual Validator
- Web Disability Simulator

Manual Testing

Screen reader

Keyboard Navigation

Color contrast

Scaling

Tool Accelerators

Experienced Testers

- Screen Readers
- Screen Magnifiers
- Simulators
- Semi-Automated Tools
- Accessibility Bookmarklets
- Open-source Tools

30% Coverage*   **30% Coverage***   **40% Coverage***

*Typically, achievable test coverage of WCAG criteria, in total 100%

WebAIM
web accessibility in mind

accenture

3. Difficulties of Automated Scorings

Automated accessibility test data and results do not always align with end user impact

What does a 100% automated test score or an "A" grade mean?

Building the **most inaccessible site** possible with a **perfect Lighthouse score**
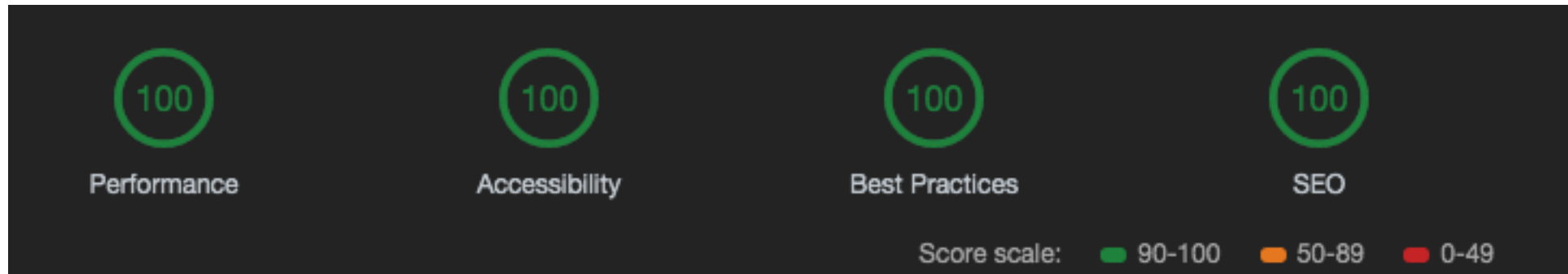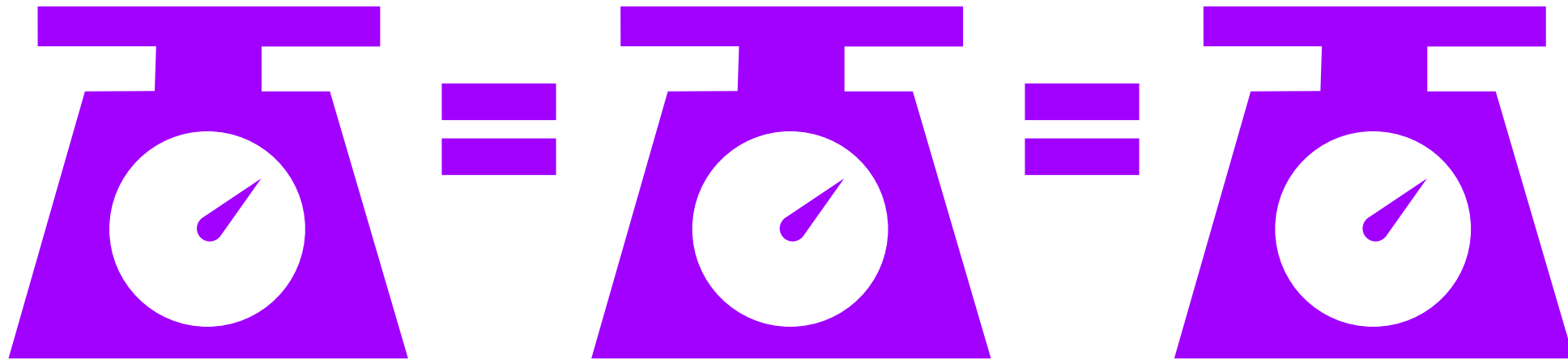
https://www.matuzo.at/blog/building-the-most-inaccessible-site-possible-with-a-perfect-lighthouse-score/

Assigning impact to automated data is often arbitrary or it favors specific disability types

1 x WCAG 2.1.1 failure
(e. g. Keyboard – Level A)

1 x WCAG 3.1.1 failure
(e. g. Language of Page – Level A)

100 x WCAG 4.1.1 failures
(e. g. Parsing – Level A)

A typical home page has about 51 automatically detectable accessibility issues

Number of issues...

... Error density...

... Content value

The typical home page has **51 automatically detectable** accessibility **issues**

(Source: The WebAIM Million)

The error density problem – to improve the accessibility score it may be easier to make the **page bigger** and **more complex rather** than **fixing** accessibility **issues**
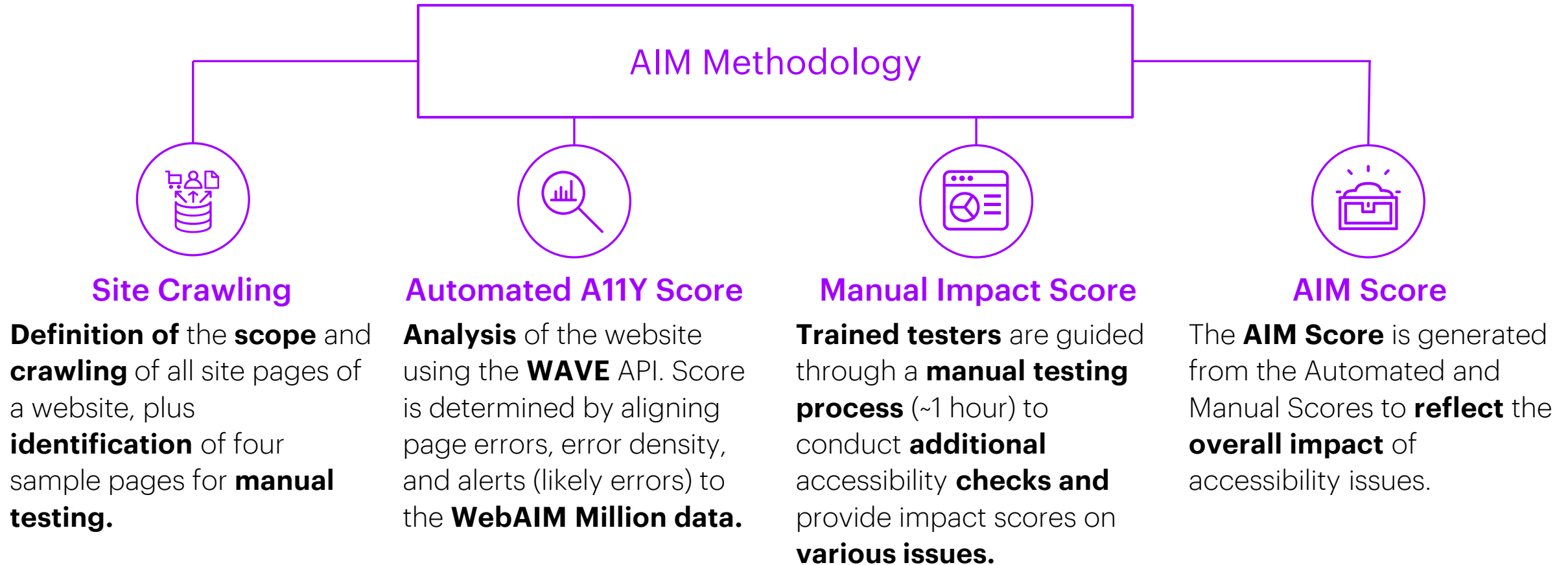
Is it possible to **factor** the **page value** or **content vs**. detected **issues**?

Manual testing solves most of these **difficulties**, **BUG** it's **very time consuming** and **expensive**

So, you want an accessibility score? – Tenon.io

4. AIM Scoring Methodology

The Accessibility IMpact (AIM) Scoring Methodology

AIM Methodology

Site Crawling

Definition of the **scope** and **crawling** of all site pages of a website, plus **identification** of four sample pages for **manual testing.**

Automated A11Y Score

Analysis of the website using the **WAVE** API. Score is determined by aligning page errors, error density, and alerts (likely errors) to the **WebAIM Million data.**

Manual Impact Score

Trained testers are guided through a **manual testing process** (~1 hour) to conduct **additional** accessibility **checks and** provide impact scores on **various issues.**

AIM Score

The **AIM Score** is generated from the Automated and Manual Scores to **reflect** the **overall impact** of accessibility issues.

AIM Score and Report Sample

For the manual testing questionnaire, we have identified the most impactful and readily testable criteria

1. Accuracy of the **document's** defined **language**

2. Impact of missing, poor, and appropriate **alternative text**

3. Impact of **empty links** and buttons

4. Impact of labeled or unlabeled **form inputs**

5. Impact of **low contrast** content (including non-text contrast)

6. Accuracy and brevity of **page title**

7. Movement and **animations**

8. Presence and visibility of **keyboard focus** indicators

9. Impact of **keyboard** accessibility **barriers**

10. Support for **page reflow**/responsiveness

Tester also records an **overall page accessibility impact score**

The Automated WAVE Scoring covers 11 out of 13 guidelines, but not all success criteria – the extent of user impact is not fully known

Perceivable	Operable	Understandable	Robust
Text Alternatives	Keyboard Accessible	Readable	Compatible
Time-Based Media	Enough Time	Predictable	
Adaptable	Seizures and Physical Reactions	Input Assistance	
Distinguishable	Navigable		
	Input Modalities		

■ WAVE Coverage   ■ AIM Methodology (addtl. depths)   ☐ Not Covered

WebAIM
web accessibility in mind

accenture

The Manual Impact Score extends the WAVE results with deeper coverage and focus on end user impact

Perceivable	Operable	Understandable	Robust
Text Alternatives	Keyboard Accessible	Readable	Compatible
Time-Based Media	Enough Time	Predictable	
Adaptable	Seizures and Physical Reactions	Input Assistance	
Distinguishable	Navigable		
	Input Modalities		

■ WAVE Coverage ■ AIM Methodology (addtl. depths) ▢ Not Covered

WebAIM
web accessibility in mind

accenture

5. Findings & Conclusions

We have applied the AIM Methodology in first practical pilots where we performed a finetuning of the scoring

Accessibility Index Report

- Sample of **30 large European web sites**
- Manual testers from **WebAIM** and **Accenture**
- All testers **rated** their sites **better** on average **than automatic** scores
- **High ICC** (Intraclass Correlation Coefficient), which adds great credibility to the manual testing process
- **High** levels of **inter-rater reliability**

Others coming soon

Johns Hopkins University

- **WebAIM** in collaboration with **Johns Hopkins** University
- **Vaccine Website** Accessibility Dashboard – tests of **56** federal, state, and territory vaccine web **sites**
- University **Disability Inclusion Dashboard** – analysis of **top 50** NIH-funded **universities**
- **Supplemental Nutrition Assistance Program** (SNAP) benefits web site accessibility rankings

WebAIM
web accessibility in mind

accenture

We are advancing this methodology through practical applications and by trying to answer the right questions

Overall conclusions:

- The AIM methodology provides a **useful (*though admittedly incomplete*) measure** of end user **accessibility impact** with **minimal costs** and **effort**

- Implementations have been **very successful** and **informative,** and provide great value to entities wishing to **improve** their **accessibility**

- A **larger sample size** and more **practical pilots** are **necessary** to be diagnostically conclusive

- Can the methodology be **expanded** to provide **weightings for error types** or by **WCAG criteria**?

- What impact will **future WCAG versions** (e.g. 3.0) have on accessibility scoring approaches?

- Can **error data** and limited manual test data be **used to extrapolate broader** accessibility **issues**?

- How can **this data** better **effect** accessibility **change**?

Questions?

Jared Smith
(WebAIM, Associate Director)

Christoph Rump
(Accenture, Test Engineering Manager)