

February 25, 2021

Accessibility Guidelines Working Group

Via email : public-agwg-comments@w3.org

SUBJECT: Comments on First Public Working Draft of W3C Accessibility Guidelines (WCAG) 3.0

To Whom It May Concern:

ITI is writing this letter to submit comments on W3C's Accessibility Guidelines Working Group (AG WG) First Public Working Draft of W3C Accessibility Guidelines (WCAG) 3.0. The Information Technology Industry Council (ITI) represents over 75 of the world's leading information and communications technology (ICT) companies from all corners of the technology sector, including hardware, software, digital services, semiconductor, network equipment, and internet, as well as "technology-enabled" companies that rely on ICT to transform their businesses.

ITI members include world leaders in accessibility who have been recognized globally for their work. Numerous ITI member companies have been very involved in the development of WCAG 2.0/2.1 and with the work of the Silver Task Force leading to this draft, including being co-chairs or co-facilitators within the time period these were developed. Some ITI member companies are also involved in the development of WCAG 3 and may be submitting comments directly; the comments below represent common feedback from all ITI members.

We appreciate the direction being taken with the evolution of WCAG and accessibility guidelines generally, as represented by this First Public Working Draft of WCAG 3. Specifically, we want to commend AG WG and this FPWD for:

- Moving away from the zero-defect definition of Conformance in WCAG 2.x. We welcome the introduction of Guidelines that recognize that there may be accessibility issues that don't have a material impact on the ability of the user to interact with a given web page or view. It allows a website or product to make a conformance claim that reflects the real-world experience and usability of a site or product.
- The emphasis on making the guidelines easier to understand and consume. Becoming an effective practitioner of accessibility in general, and web accessibility in particular, requires a tremendous amount of specialized technical knowledge, as well as a deep understanding of user needs of people with disabilities. There is still work needed to make the standard easy to use, and this is particularly true for people with long experience with WCAG 2.x because of the new concepts and terminology introduced in WCAG 3.0, especially related to the conformance model. We agree with the goals around making the guidelines easier to understand and expect that this will be incredibly helpful to the large number of website and software developers who

may be tasked with ensuring the websites they publish and the products they build are accessible to people with disabilities.

- Updating guidance in response to current research. For example, the draft Guideline 7.5 Visual Contrast of Text (<https://w3c.github.io/silver/guidelines/#visual-contrast-of-text>), which more accurately models current research in human visual perception of contrast and light intensity, is in line with ITI member companies' findings in user studies on text contrast.
- Support for more users with disabilities. We agree that it is important to provide additional guidance for users with cognitive disabilities and other users, but also recognize that this can increase the overall level of effort required to implement accessibility support, which supports the need for improved testing and enhanced conformance requirements.

Also, we want to raise with you several topics that we believe must be addressed in order for WCAG 3 to be effectively adopted and used. These include:

- Outcome ratings vs. critical errors vs. conformance levels is confusing. Adding to the complexity is the objective pass/fail outcome vs adjectival ratings, making it even harder to follow. Understanding and meeting the outcomes is hard enough for people to learn that we don't want to see there be a complex scoring added on top.
- ITI believes that in order for organizations to embrace using WCAG, the FPWD scoring information must be more clearly articulated in order to explain it to all readers in an effective manner. This could be achieved by providing more examples, models, and detailed explanations.
- ITI appreciates that the example in 6.2 for an adjectival rating scale clearly states how things should be measured, however it appears to be an almost impossible bar to meet. If there is an expectation to achieve a 3.5 point overall score, based on the example listed, that would translate to the requirement of a 95% passing level on all tests. Further, the table in 6.2 sets the bar too high (95%). We appreciate the Editors Note that the number could be 90%, which may be more reasonable, underscoring the need for the entire scoring mechanism to be much more clearly stated.
- As WCAG 3.0 does not supersede WCAG 2, there needs to be a mapping of the success criteria from WCAG 2 to WCAG 3.0 outcomes. This will facilitate both product creators, with their development/testing/reporting of products previously evaluated using WCAG 2, and product consumers so they can use their knowledge of WCAG 2 to bring them up to speed on guidelines in WCAG 3.0.
- Guidelines need to be testable/measurable based on normative information. In some new guidelines such as 7.2 Clear Words this does not appear to be testable based on normative information. In the Methods the tool Simplewriter is listed. There is no attribution for this tool (who owns it?) Instead of recommending a specific tool, WCAG should specify the criteria that should be tested, so others can incorporate it into their current tools or look for a publicly available tool.

Finally, we have identified a number of topics that we believe are critically important to be addressed in the next published Working Draft of WCAG 3.0:

- Inclusion of a compliance level or mechanism that would accept a very small minority of pages having one or more "critical errors" on them, while still achieving a minimum compliance bar. Doing so would align WCAG with the real-world tradeoffs that businesses have to make every

day with a balance of time-to-market with similarly critical non-accessibility bugs that may prevent a specific feature from being used by all users, or by users using certain devices, or from certain geographies, or some other subset – which are then fixed quickly, post-release/publish. We believe this is particularly important for the very largest and most dynamic of websites and web applications, which are undergoing continuous updates, and which at any given moment in time almost certainly have some number of significant issues appearing somewhere on their site, somewhere within the application.

- Inclusion of a compliance level or mechanism that only depends upon automated/automatable tests rather than needing human judgement and review. Such a compliance level or mechanism might be an ever-rising bar, that ratchets up over time as machine learning and other techniques allow more comprehensive automated testing. We believe this is particularly important for the very largest and most dynamic of websites and web applications, such as those driven primarily by user-generated content, and which may be receiving more than 100m updates per day (more than 50k updates per second).
- Explicit treatment of, and allowance(s) for, a broad range of 3rd party content including:
 - Content not owned by the website owner (e.g., movies, ads, maps, etc.) where copyright restrictions or license agreements may prevent the website owner from changing the content to make it accessible.
 - Content shared by individual website visitors (e.g. comments, blog posts, social media posts, recorded media). While a website may offer authors (visitors) the ability to submit accessible content, we are concerned with the impact of holding sites accountable by requiring visitors to submit content through an accessibility assessment and remediation pass (e.g., having to put ALT text on every image they contribute; having to meet language simplicity rules for readers with print/cognitive impairments). Doing so places a burden of specialized knowledge on the part of content contributors, and so could result in contributors “gaming the system”, such as adding garbage ALT text to pass the content posting requirement. It could also lead to a large drop in content sharing.

In addition to these comments, ITI is also providing responses to specific questions posed by the Working Group as an attachment to this letter.

Thank you for your time and consideration of our comments.

Sincerely,



Erica Thomas

Questions posed in the WCAG 3.0 materials and draft

Questions from the blog post [WCAG 3 FPWD Published, Feedback section](#)

1. Are there additional Design Principles and Requirements that should be included in the WCAG 3 project?

Yes - inclusion of a compliance level or mechanism that only depends upon automated/automatable tests rather than needing human judgement and review. Sites need to be able to approach some level of conformance without manual effort, or with minimal manual effort on a limited number of key outcomes. This would, of course, not be a desired end-point for any content, but this is sufficiently important for many large and highly dynamic sites that it merits mention in the requirements document.

2. There is considerable introductory information, including information about the guidelines structure, testing, and scoring. Are there usability improvements that would make it easier to use and find information?

Yes - the guidelines structure should be in an explainer document and this document focus on the requirements. Testing and scoring should provide better practical usage with examples, as it is hard to comprehend the bigger picture that leads to the overall conformance and score.

3. Outcomes are normative. Should guidelines, methods, critical errors, and outcome ratings be normative or informative, and why?

Guidelines should be normative, and normative text should also indicate that conformance with the guidelines is determined through aggregate scores for outcomes. The outcome ratings and critical errors also need to be normative as these are used to determine overall conformance with the standard. Methods are technology and situation-specific, much like the current WCAG techniques, and these should be informative.

4. We would like constructive feedback on the testing approach, and examples of why you would or would not implement it in your organization.

The testing approach with adjectival ratings tied to more detailed summary evaluation seems worth pursuing, but the thresholds for the ratings need to be clearly linked to some metric that helps clarify the logic between the rating level and the requirement for overall conformance to bronze, silver, or gold.

5. Is this approach of using complete processes as the smallest unit of conformance workable for different types of organizations? These include organizations with very large, dynamic, or complex content, medium-sized organizations relying on external accessibility resources, and very small organizations with limited resources?

It was not immediately clear that the smallest unit of conformance is a process, so the group should review the document to make this more clear, including adding this to the "Differences from WCAG 2.x" section.

The use of process as the unit of conformance, which could include a single page in some cases, seems like a useful concept. The challenge with this is in the definition of process, which is rather loosely defined as "A process is a complete activity the user performs.". There may be a challenge when content providers - disagree with others on what constitutes a process. Having examples of what is not a process would help greatly.

6. Does the model for scoring and aggregated ratings work? Why or why not? If not, please propose an alternative solution.

We respond to this in the Conformance Model questions below.

7. As you evaluate this document, please consider whether there are ways the Working Group can better support your review, feedback, or inclusion within the process of creating this standard.

For future drafts, offering a consolidated list of questions will help facilitate review and comment. Currently questions are provided in multiple locations and are sometimes repetitive.

Questions in the WCAG 3.0 FPWD Status of this document section

1. How well will this proposal of conformance work in your type of organization and why?
2. Conformance level approach presented in the document
3. What do you see as the possible tradeoffs in using a flexible conformance approach? Would you want to see multiple conformance models?
A flexible conformance approach might be a vehicle for addressing the more complicated issues involved in 3rd party content conformance, as well as flexibility in describing the main/primary functions of a web page or view vs. the secondary or tertiary ones. However, multiple conformance models would likely make conformance claims more complicated and confusing. There should be one model for conformance, based on rating and scoped by documenting the sampling, processes, and any other testing performed with an easy way to calculate the conformance level. If it is too complicated, it will be difficult for policy makers to understand what conformance levels can be reasonably expected for technologies to attain that can be included in regulations. More clarity in the specific proposals is needed to give a more detailed response to the question.
4. How will the change to a different conformance model than WCAG 2 affect your organization? How readily would you see moving from WCAG 2 to WCAG 3 when it is complete?

5. Are there usability improvements that would make WCAG 3 easier to use and find information? Is the structuring of the How-To, outcomes, and methods clear? Are there improvements to structure and style that you would like to see? Is there another way we can make content more usable?

Questions from the body of the main document

Section titles are named on each item so you can find the referenced topic in the document.

1. The working group is looking for feedback on whether the following should be normative or informative: guidelines, methods, critical errors, and outcome ratings.
2. **5.3 Types of tests:** We are looking for more appropriate terms to distinguish between these two types of tests and welcome suggestions.

ITI agrees that the names for the two types of tests referenced in 5.3 are likely to cause confusion. In particular, "Atomic testing" appears to be in conflict with the definition as used by the ACT Task Force. One possible suggestion for replacement would be to replace "atomic testing" and "holistic testing" with "Module testing" and "end-to-end testing", respectively.

3. **5.4 Technology specific testing:** We welcome suggestions on ways to improve the scoring to better meet technology-specific criteria.
4. **6.3 Overall scores:** Invite comment on carrying percentages from tests through to a final score OR ratings are averaged for a total score and a score by the functional categories they support. E.g. bronze conformance level means no critical errors and least 3.5 total score and at least a 3.5 score within each functional category.
5. **Conveys information about the sound:** looking for feedback on scoring of captions and the approach of scoring essential videos as more important/higher weight than captions on advertising and promotional videos. Cumulative lack of captioning further reduces the score. Are burned-in captions equivalent to closed captions or should they be scored lower?
6. **Conveys hierarchy with semantic structure:** We are looking for feedback on using scoring as a way to encourage adoption of AAA success criteria without failures. Do you like the inclusion of broader needs for structured content than providing semantics for screen reader users? Do you think this should be a separate guideline, or do you like having multiple, testable outcomes supporting the guideline? Do you like the approach of merging WCAG2 success criteria with related user needs?
7. **8.0 Conformance model:**
 - We seek feedback on the conformance approach as a whole.

Regarding the current state of the conformance model, there is insufficient detail to fully evaluate it. The general concept of defining thresholds for adjectival ratings and scores is good conceptually, but several questions remain:

- I. How are the thresholds for the adjectival ratings determined?
- II. Will the thresholds be stable over time or will they change to become more strict or possible to become more lenient as changes in technology provide improvements to support end users (or will those be reflected in the outcomes themselves?)
- III. The threshold of 3.5 for Bronze conformance represents another potentially adjustable value that impacts conformance. The draft indicates that the intent is to align bronze conformance with a 90% “score” but it is very difficult to see how that alignment is achieved when looking at the ratings for the outcomes that are currently presented. We believe that the general approach is workable, but wonder why the additional 3.5 factor is needed – couldn’t the outcome ratings be adjusted to eliminate the need for this multiplier? Can the group provide a rationale for why the multiplier is needed?

It would help evaluate and understand the scoring scheme if there was a sample scoring sheet provided. It is very hard to determine the difficulty of conformance scoring across different functional groupings and in aggregate. Understanding whether the model is workable without the ability to view it in this context is a substantial challenge.

Right now, the only level specified is Bronze, as the holistic tests are not defined for Silver and Gold levels. While we cannot comment on the holistic tests, a key concern with these tests is the level of effort required to conduct the testing and how objectively the requirements can be assessed.

We do note and appreciate the Editor’s note on addressing representative sampling in a future draft. This is a very important topic, and relates to our point about flexibility for conformance.

- We seek feedback on whether the flexibility introduced will encourage content providers to meet conformance because it is more achievable vs. whether they are less likely to improve if they don’t have to.

Flexibility for conformance is desperately needed. Full conformance is very difficult for most sites and products. Full conformance is impossible to evaluate for large and highly dynamic websites, and complex software. Flexibility in the conformance model will provide a substantial incentive for adoption, although it is unlikely that sites will opt to ignore WCAG 2.x conformance efforts. A mapping between the WCAG 2.x and 3 standards would also encourage adoption with a clear path to move to the new standard.

- Representative sampling is an important strategy to large and complex sites. We welcome suggestions and feedback about the best way to incorporate representative sampling into WCAG 3.0.

Representative sampling will be a highly customized process for different content providers. Sampling techniques have different issues for different types of sites. For example, a sampling process that requires manual testing on 5% of views would be a

major challenge for a site with 50,000 views (2,500 manual view tests), let alone one with millions. A requirement that the content provider selects the top ten workflows to test is much more achievable. However, workflow definition may vary greatly between providers so guidance is needed on what level of detail should be included.