# Review of WCAG 3.0 first public draft

Gregg Vanderheiden 01-March-2021

This review is divided into five sections

* Laudits
* Concerns
* Problems
* Applying these comments to the three provisions provided as examples
* Recommendation/possible solutions

The first half of this review - lays out the concerns and issues.

The second half - applies them to the current guidelines and talks about solutions.

I also have editorial comments but will submit those separately as they are of a different character, easily solved, and usually just the result of oversights in trying to get the draft out to the public.

## Laudits

### 1) Taking a holistic view of accessibility.

It is good to see a document that looks at accessibility as being a collaborative effort between content creators, user agents, and tools. This approach can cause some problems around conformance, but I think there is a way around those as described at the end. I understand how difficult and complicated task it will be to undertake if you try to do anything besides just describe status quo. Nonetheless, it is a very valuable contribution and kudos for attempting into. I also recognize that doing it right will considerably slow the process of developing a 3.0.

### 2) Better coverage of cognitive language and learning.

The document also seeks to find a format that will allow better coverage of cognitive language and learning disabilities. To the extent that this document can do that, then it (along with the new W3C report specifically on cognitive language and learning disabilities) can provide greatly needed contributions to understanding these disabilities and documenting different strategies for addressing those of us experiencing them.

### 3) Outcome orientation

The document describes itself as being outcome oriented. This is excellent. This is what WCAG 2.0 also focused on. That is, it focused on “what should be achieved” but not “how to achieve it” (except for the techniques doc of course). This (outcome) approach allows for the greatest interplay between content developers and user agents, dividing the load between them and allowing for advances in user agents to remove requirements on authors - without rewriting guidelines. (For example, a criterion that says that text must be “able to be enlarged” by some factor talks about an outcome not an imperative. As a result – when browsers built zooming (to that level) into their browsers, authors no longer had to do anything except not break zoom functionality of the browser. In the future the zoom functionality of the browsers may be even more powerful, further reducing, and perhaps someday eliminating, the need for authors to do anything in order to have pages that will re-layout and reflow as users zoom them.

Thus, keeping the provisions outcome-oriented is a way of both providing authors with the greatest flexibility in terms of meeting the provisions, but also making the provisions more future proof, or at least future resilient.

Making ***outcome-oriented*** provisions, however, takes a lot of thinking and work. It is noted here that even with the small sample of 3 provisions provided, only two are***outcome-oriented***. It is very easy to slip into the ***imperative*** or ***command*** or ***directive*** format that dictates “what authors should do” rather than “what should be achieved”

## Concerns

### 1) Model-based guidance (top down without simultaneous bottoms up)

I was going to put the Model Based approach in the ***Laudits*** category above. It's good to have a theory or model behind what you do.

However, all models need to be also data-based rather than just theory-based to be useful. That is, they need to model “what has to happen” and “what can happen” not “what we wish was possible”. Thus, good model building involves both a top-down and a simultaneous bottom up look at the topic. A bottom-up approach requires that you first gather all of the things that you are trying to model or organize and study them before you start trying to draw top-down.

You can't design an organization, or a structure, or a model for something until you actually first look at all of the parts of what you trying to model. In this case, the group should have a full set of the provisions for the new standard – or as many as possible. Not only the old provisions but also all of the new provisions on the table that they are trying to create a model for.

If this had been done, (and the rules for qualifying a provision were applied) some of the problems cited below may have been avoided. As it is, the document spends all of its time laying out it's elaborated model for provisions and testing etc. without having the provisions and tests in hand.

As a result, everything may seem to fit a model when in fact it does not when you actually put things in it. In some ways it reminds me of Marie Kondo who says don't organize your closet until you first empty it all out, sort it and see what you have. Otherwise, you end up with an organization that doesn't actually fit the clothes that you have in number quantity and type. In this case we are not talking about sorting, (it would be nice if it was such a simple task) but something much more complicated. The underlying principle however is the same in all of science. Models that are created without data are of limited value in practice. Models that are not tested by applying them against data are speculative.

It is also important to remember the quote by George Box.

***“All models are wrong, but some are useful”.***

That is, a model can be useful in guiding your work, but adherence to the model often causes you to force things to fit that don't really fit. In WCAG 2.0 we also created models for provisions and, for example, what would be in level A, AA, and AAA. However, when we actually came to assigning provisions to the different levels using our model/rules- we found that the model/rules didn't work. We kept adding more and more factors, but it still never worked. In the end we found that the data (the provisions) and the reality of the needs of different groups, as well as the interaction between different provisions informed the decisions. As a result, the answer to the question "*what were the factors determining which level something when into*" ended up being a list of factors that were used, but no strict formula -- because no strict formula could account for the complexity.

Similarly, the models in WCAG 3.0 may be good for thinking – but are unlikely to work as planned when you actually try to apply them to all of the provision – and the constraints in creating provisions. Suggest that, with this modeling phase being done – you focus next on collecting all of the provisions and seeing if they fit your model – or you need to rethink, revise.

#### Top-down and Bottom-up in combination.

Actually the best approach usually is to do both top-down and bottom-up together. Since you already have WCAG and the work of the COGA group – you have a first pass at bottom up. You now also have a first pass at top-down. The next step is probably to gather all of your potential provisions (from WCAG and COGA and other disability efforts, then look at them and come up with your first-final set of guidelines. (if you want a document that can be used in regulatory work – you would have to run them all through a filter to get a set of objective, testable, items. If going the guidance route – you can include them all.) You can then take these and fit them into your model to see if it works. If not – you can modify the model as needed. And the iterate on provisions, model, and the language / wording for both.

### 2) Testing

Two concerns arise around the use of the word testing in WCAG 3 draft.

1. Things seem to be grouped in testing that are not actually tests per se
2. Testing for conformance to a standard is conflated with testing to make sure a site is as accessible and usable as possible. These are two very different activities. Both important. But not that same.

In some places Testing is used to describe something which is a real test. That is; a test to be sure that something does something? Does something meet a criterion?

In other places (i.e. in the section on holistic tests) it says that testing includes ‘user-centered design methods’ and ‘usability testing’.

* **‘User-centered design methods’** is not a test. Testing may be used as part of the user-centered design process, but user-centered design is not a test or a method for testing.
* **Usability testing** ***is*** a means for testing, but not for conformance testing.
	+ Usability testing is used for continuous quality improvement (CQI). These guidelines for example should be user tested to make sure that they are usable. However, they can still be an accurate standard of the highest quality, and not be usable by many people who just don't understand standards, or the technologies needed. You could also create guidelines that are very usable by people (for example by making them very short and simple) but have no validity or real use in testing whether something is accessible.

Tests for conformance to a standard need to be restricted to tests that actually test conformance to the standard. I agree that ALL websites should be user tested - by people who have a full range of disabilities. But that is a completely different issue from testing the sites to see whether or not they conform to a standard.

* Conformance testing determines if a website conforms to a standards.
* User testing is done to improve the accessibility of the site and to identify issues that were overlooked or not covered by provisions in the standard – but still important.

My concern is that the term ***testing*** is being redefined in different ways than normal for conformance testing - in order to try to achieve some goal that is outside the reach of an accessibility standard. For example, if one couldn't find any ***objectively testable provisions*** that would cover the needs of a particular disability group --- but one really want to include more provisions for that group…. one approach would be to re-define what is meant by ***testing*** to include different types of ***nonobjective testing***. This would then allow one to include ***nonobjective provisions*** – which in turn allows more provisions for this group.” That is fine for a general guidance document – but not for a standard that one needs to conform to.

I know the pain that can lead one to do this. We had the same problem in WCAG 2.0. We actually spent enormous time on cognitive language learning disabilities for example (more than on any other single disability) trying to find provisions that would address their needs and yet would be objective and meet the criteria necessary for a testable provision. We called in Nancy Ward, Clayton Lewis and a whole host of other people to talk with us and propose provisions that might work. John Slaton and I launched two, many-months-long efforts on both the cognitive language and learning disability area and the use of plain language in the guidelines. It was the most frustrating thing I have ever done in my life. Seeing the needs, but being unable to identify or find ways to qualify as strategies from all the materials we read, and people we talk to, was the most difficult and frustrating part of the work on WCAG.

### 3) Mixing of three sectors in one guideline

There is a slight concern regarding mixing guidelines for three audiences in one document. Will it confuse each audience with regard to which provisions apply to them. But this is easily remedied by either organizing the document into sections – or tagging the provisions to indicate which apply to each group. I think many will apply to all groups – and having the responsibilities of the different sectors all side by side is a feature of this document. Provisions for that apply only to one audience, however, should be separate.

## Problems

### Conflicting goals

WCAG 3.0 appears to have three underlying goals.

1. To create a document which takes a holistic view of accessibility across content, agents and tools..
2. To create a set of guidelines that can include provisions for people with disabilities that goes beyond what is allowed using the WCAG 2 rules for inclusion (e.g objective, testable provisions etc.)
3. To create a document in the WCAG 2.0 line that will be picked up by the field and regulatory agencies and thus given the weight and power that a standard or regulation would have. (Hence the efforts to make the initials of the group come out to be WCAG, and its number to be 3.0 even though is the actually the first in its line.)

All of these are needed and laudable (and lauded above). But goals 2 and 3 cannot be met by the same document.

In order to be a standard that is picked up in regulation - it has to be objective, and all of the tests and provisions need to be objective. One simply cannot meet this goal and still have the type of document described in (2) above.

The authors seem to be aware of the inability to address #2 and still have objective criteria and tests. Great pains have been taken in WCAG 3 to describe all sorts of testing that is non-objective (e.g. the use of qualitative rating scales, qualitative rather than quantitative or binary measures, usability testing, etc.) in order to be able to achieve goal two. Unfortunately, this disqualifies it from goal three.

In the end the group will need to either rename the document and have it be a really wonderful guidance document with broad scope for including guidance provisions, or return to the WCAG 2x like criteria for selecting provisions that is needed in a standard that could be adopted in regulation. This latter choice would, of course, put you back in the same bind as the existing WCAG 2 thread. It is aggravating, bang-head-against-the-wall frustrating, etc. but that is the situation. (See more detail below)

### Introduction of non-objective tests in WCAG-3.

In order for something to be a standard, particularly a standard that is going to be used in regulation of any type,

* all of the provisions that are normative (that is, the only ones you have to pay attention to in order to conform) must be objectively testable. That is, one must either pass them or fail them.
* And you must have high inter-rater-reliability (that is, if you have a number of people who are aware of the technologies in use, a very high percentage of them would all come up with the same answer as to whether something passes or fails.
* If a number is used, (e.g. the volume must be above 3 dB), then the scale, the number that is the threshold, and the tool that can be used to measure the threshold value (number) that causes it to pass or fail, needs to be cited and the instrument creating the number needs to be objective.
	+ For example “The background noise must be 20 dB down from the foreground speech" would qualify.
		- but “The evaluator's opinion is that the score is a three or better" does not qualify since it relies on the opinion of an individual rather than on a measurement.
	+ Any measurements that rely on opinion are called *qualitative* rather than *quantitative* and are *subjective* rather than *objective* measures (unless one is measuring opinions rather than conformance).

This constrains the types of provisions or requirements that you can have in a standard. Often leaving out guidance you would like to include but cannot reduce to an objectively testable requirement.

We ran into the same problem repeatedly in the WCAG 2 series. For example: we were able to require that the text alternative be available for any picture- since it is easy to objectively determine if a text alternative was present. But we could think of no objective measure that would allow us to require that the text alternative be of any particular quality. We tried length and many other measures of alt text quality, and they all failed. In the end we were only able to require that there was alt text for any (non-decorative) image. We pushed it a bit in our “common failures” list for the provision by saying that the *filename for the image* was not sufficient if inserted as alt text. But that was advisory not normative.

**Rationale / Why is it so important that everything be objective?**

First, it is unfair to an author to ***require*** that they do something, if they have no mechanism to determine if and when they have done enough. For example, requiring that something have “sufficient contrast" is not good enough since an author would have no idea what contrast was sufficient. Similarly requiring that they use clear language is not fair to the author unless it is objectively stated exactly what “clear language” is and exactly how to objectively measure it – along with a clear threshold for passing. Clear to whom? How clear? How do I measure "clarity"?

The second reason this is important is because regulations are completely useless if they cannot be enforced. And the only way that a regulation can be enforced is in a court of law. Regulations stem from laws and laws are enforced by the judicial system. (Occasionally there are other bodies designated such as the FCC or mediation bodies etc. In all these cases someone still ends up "judging" whether or not the provision is met.) If the provisions in a standard are not objective, the very first time it shows up in court, the defendants will cite, accurately, that the provision is not objective but rather is subjective. And as a result, it is not enforceable. That is a real danger here. The WCAG 2 series has so far not been challenged as being arbitrary. The major problems we have had to date have been either around whether the law actually was meant to cover the web (which although it's been decided several times keeps being challenged) or around the cost for making things accessible (which we all know is higher if you build everything first and then try to retrofit accessibility into it).

It is doubtful that a nonobjective document in the WCAG series would be able to make it through to a “consensus document without objection” (as WCAG 2 was). And if it was adopted over objections by the W3C, it is doubtful that international standards bodies would adopt it as they did the WCAG 2 series. Introducing something with the WCAG initials that is not objective would thus not only *not* achieve the objective of following along in the path of WCAG 2, but could hurt the WCAG line itself as well as cause confusion in the field. There has already been pressure from and on other standards bodies to shift from WCAG and the W3C for setting web standards to ISO or other international standards bodies to create accessibility standards. Up to the present it has been possible, based on the strength and quality of WCAG 2 series, and the backing of industry, to have the other international standards simply refer to WCAG for their web content accessibility provisions. Creating a WCAG 3.0 that does not meet the requirements and standards of these standards bodies for provisions in their international standards could cause them to break from using the WCAG series in the future. This would not be an arbitrary or capricious decision. They simply would not be allowed, by the rules in their standards organizations, to include normative provisions that were not objective. You will also note, that in their standard, they do not even allow three levels of conformance. In their standards you either conform at level AA or you fail to conform. For example, you do not have a choice in EN 301 549 to conform at either level A or AA. Both A and AA provisions are included in their standard – and the requirement is that you conform at level AA or you don't conform.

### Complicated conformance and scoring methods.

It is fine to have a complicated scoring or rating methods for a guidance document (e.g. one that is not called WCAG and which is not intended to be adopted in regulation). However, if one wants to create a document that would be used in the legal and regulatory arena, then one has to stick with objective measures.

It should be noted that even when creating WCAG 2.0, where guidelines were first sorted into 3 levels (A, AA, and AAA), it was always understood that we were simply giving regulatory bodies a choice of choosing either the provisions on Level A or the provisions in Levels A and AA (as a group), as the provisions for their standard. There was never any belief they were going to have two levels of conformance in their standard. We simply gave them two choices as to where they wanted to draw the line for accepting provisions into their standard.

It should also be also noted that we also created a level AAA that contain provisions that we tried but just could not get into level AA because they did not meet the criteria. We stretched the requirements here in creating Level AAA because wanted those concepts and needs to be in the same document -- so that they would be able to get the attention of everyone -- and so that we could write techniques and strategies about them. We also added lots of techniques and strategies that went beyond what was required in level and AA and attached them as additional guidance for the same reason. I see a similar strategy being tried in WCAG 3 – but not identified so clearly. In WCAG 2 we clearly stated that Level AAA was different than A and AA and that it should not be required across all content because they did not meet the tests for what could be required for a provision that would be objective and/or that could apply across all web content.

I mention these here not in defense or justification of WCAG 2.0, which needs to stand on its own and be critiqued and improved as was done in 2.1. and 2.2. I mention these here so that there is a better understanding of the frustrations and the barriers to be faced in the WCAG 3 work – and which, as I highlight here, were also faced before, and how we went about trying to deal with them. Some “missing” issues were not covered in earlier versions not because those WCAG teams didn't care, and not because there were not people in those working groups who were interested in and worked hard for all of these disability groups and issues, -- but simply because they/we could not find ways to cover these issues… could not find language for provisions that would work.”

**"But we aren't writing a regulation – so why do we have to follow those rules/constraints?"**

Yes, this is true. This is an accessibility standard - not a regulation. We said the same thing when we were writing WCAG 2.0. But we quickly realized that if we wanted those who *were* writing regulations to use our document, that we need to conform to their rules. If we didn’t/ don't follow their rules, then they will not be able to use our document directly. They will have to make up their own provisions or pick and choose among ours to find those that do meet the rules that they need to follow, and leave all the rest behind.

Again, it is suggested that a decision be made as to whether (A) this document meant to be a really great global holistic document on accessibility, or whether (B) it's meant to be the next chapter in the WCAG line of accessibility standards that were/are designed to be able to be used in legal and regulatory spheres.

### Role and use of User Testing

* User testing is really important for continual quality improvement (CQI).
* User testing should be done on all websites and should involve a diversity of users.
* User testing should be done on these guidelines to ensure that the guidelines themselves are as usable as possible.
* User testing is important to carry out in the creation of guidelines.
* But user testing is should not be used to test for conformance to a standard.

User testing is not designed for and not a valid method for conformance testing - as would be done with regulatory standards. The reason is that user testing is not objective. It would not have high interrater reliability. That is, if you bring in 10 groups of users you are likely to come up with 10 different results (overlapping but not the same). So, having an author do user-testing will not guarantee that anything will meet some objective level of accessibility. In fact, I can assemble a group of individuals would like to use the web who likely would be able to fail any interface on the web including any interface designed by the working group. I can also bring in a group that could use websites that fail WCAG.

***Thus, user testing is very powerful and very good for what is it good at:***

* increasing the quality and usability of a site,
* looking for missed issues
* identifying issues not covered by accessibility standards
	+ in order to consider them for inclusion in a new standard – if they are objectively measurable
	+ in order to make a site more accessible than standards require (go beyond or fill in gaps)

But user testing does not have a place in testing for conformance to a standard. It is not objective and does not have the necessary inter-rater reliability needed that an author could rely on it to answer their question “does this site pass”. Nor to be used by an evaluator to prove that it does or doesn’t meet some stable, defined level of accessibility.

**User testing does have a role in standards.**

But it's in **the development** of standards **not in testing conformance** to standards. Let me illustrate with an example that is a little easier to visualize.

Suppose you are creating standards for wheelchair access to bathrooms

You would not want some engineer to measure a wheelchair and its turning radius, and then decide what the standard bathroom measurements should be for accessibility.

You would want to do a lot of user testing, where you had real people with disabilities, and real wheelchairs, navigating real bathrooms - in order to understand what works and determine what the standard would be.

Once you have the standard, you would want it to be clear and objective so that the person building the bathroom could be held accountable to it. They would knew exactly what the rules were, and could tell exactly if they had made met the criteria not. It would not be up to inspectors to decide if it looked good enough – they would be able to (and be accountable for) making sure the clear rules were followed.

If one used user testing to determine if the bathroom was accessible

* Builders might choose one set of users – and find they could use a bathroom
* Only to find the inspector used a different set of users – and came to a different conclusion.
* Different buildings would have different dimensions depending on which users were used to test and pass the bathrooms.
* And if it ever ends up in court – (the only way to enforce regulations) then any testing based on individuals would fail as being non-objective and arbitrary based on which users are used. And the court decides what is reasonable rather than the standard.

So user testing is not good for testing for conformance to a standard.

However, as noted above, it IS good when developing a standard to help identify issues and determine what would be a reasonably effective threshold value for any measures (nothing is 100% good enough for all people).

User testing is also very important for companies to use in testing the accessibility of their sites or products – to figure out how to improve the accessibility of their sites, – to identify things they may have missed, and – to identify problems that may not have been covered (or might not have been able to be covered) by accessibility standards.

### Use of scales and ratings.

As noted above, these can be very powerful in a guidance document.

But the only scales that can be used in a standard (that is meant to be used or referenced for regulatory purposes) would have to be scales that are objective in nature (e.g. no one's opinion) and where the exact value of the scales is specified where pass/fail occurs.

This means that any scales like good/better/best, or partial, or things like this, are all good for guidance and for measuring the progress that a company is making in making their websites more accessible. However, they cannot be used in something that is objective. That has conformance.

### The use of the word “users” within a provision

This is a problem we have faced in all of the standards groups that I have worked on dealing with accessibility. We found that we continually had a tendency to inject the word "users" or “users can” into our normative provisions. This only caused us to later have to go back and rework them to get them back out.

For example, any normative statement that says that something “should be done in order for users to” runs into the problem of not defining which users one is talking about. For example, “Alt text should be provided that will allow users to understand the meaning of the image”. Which users? We cannot be talking about *every* user in the world, because there will always be users who are going to be completely unable to use whatever technology were talking about.

For example, diagrams on a website on quantum physics or string theory are not going to be used or understood by a large number (most) users no matter what alt text you provide – because they couldn’t understand it if they could see the diagrams. The same thing about advanced cooking, or woodworking or a lot of other things. There will always be websites that are talking about things that are more advanced and require more advanced knowledge than a user has. There's nothing one can do to that website that is going to make it understandable to that individual. Even advanced knitting sites.

But the same goes for shopping sites. There are some (a much smaller number) who will be confused by any shopping site – and would be confused if you left them unattended in any physical store as well.

We can’t mean to require that all users will be able to understand. You also can't define “users” as people who are already using it, or else that's a self-defeating requirement. You can define it as “all those who ‘could’ “ since you have the same circular problem.

Having said this about not including “users” in the actual provisions, I will note that including them in the RATIONALE for a provision is fine and very useful.

For example

A text alternative shall be available for all non-text content.

Rationale: This allows the non-text content to be understood since the text alternative can be rendered into visual, auditory, or tactile form – or translated into different languages or symbol sets to meet the needs of individual user – just like the text portion of the website can.

## Applying these comments to the three provisions provided as examples

In this section I'm going to take the three provisions that were provided examples and talk about how the above elements of apply.

### 7.1 text alternatives



This starts out with an imperative or directive rather than an outcome.

The imperative is ***"Provide text alternatives for all non-text content***”.

This would be comparable to the WCAG 2 Guideline 1.1 text alternatives: provide text alternatives for any non-text content so that it can be changed into other forms people need such as large print braille, speech, symbols or simple language.

This is then followed in WCAG 3.0 with the outcome "text alternative available”

This is a good statement of an outcome.

However this is then followed by something which says

" Provides text alternatives for non-text content for user agents and assistive technologies. This allows users who are unable to perceive and / or understand the non-text content to determine its meaning”

This text is unlabeled (as is the Objective). If this text was labeled as **benefit** or **Rationale** than this overall 7.1 provision would be pretty good.

1. It has a general sentence above (Provide text alternatives.. ) that is not-normative but helps the user understand the objective
	1. This is similar to the WCAG 2 ‘Guidelines” which were also non-normative. I would suggest though that if reformatting – we might look to eliminating imperatives and focusing only on outcome-oriented statements (see comments above on future resilience in the Laudits section)
2. It would have a very simple outcome-oriented statement which is testable "text alternative available" – and would be normative.
	1. (It could be a bit less ambiguous since it seems to require that there be a text alternative for text. It would probably be better if it read "text alternative available for any nontext content" – or easier to read by inserting the word is “**A** text alternative **is** available for any non-text content”.
	2. This should be clearly labeled or in some way clear that it is the “Desired Outcome” or “Requirement”. It currently is only labeled as such in the markup – which requires you look at the source to discover this. Using visual formatting to indicate that this is the key phrase – violates 1.3.1 of WCAG. So, some text labelling should be provided to assist both those with visual disabilities and those who have trouble processing complex documents sort out the key/core text of the provision.
3. It would then have a Rationale that explains the purpose and benefit of the provision making it easier to understand and providing motivation to conform.
	1. This however should be labelled as such – as noted above.
	2. This would be non-normative.

Note that the comments about users being able to understand etc. are all in the rationale and so they do not trip the concerns I cited above about discussing users within actuals standards normative comments.

### 7.2 clear words



The first line of the guideline is “Use clear words” which is an imperative. As with 7.1 this is non-normative so it is ok as an imperative – but I would suggest we move away from them even in opening lines like this. In WCAG 2 they were used to open up a whole section. Having an imperative for each provision – I think is confusing and I would advise against it – and moving toward the Outcome orientation approach cited in the introduction to WCAG 3.0

The next part is an exception. This is very confusing on a number of fronts.

* First – the exception comes before the Objective – which is the normative part of the provision. Having an exception before the requirement is confusing. Perhaps this is just an editorial oversight.
* Second the wording of the exception is a problem – but I’ll discuss this below after going through the rest of the parts of this provision.

Next comes the outcome-oriented requirement “Common clear words”. However, although it is labeled in the markup as the objective – and is in the same location as the objective in 7.1 – it looks more like a title than an outcome. I would expect an outcome to be worded more like “content is available in clear common words”.

Another problem with this provision is that it is not objective unless the words ‘clear’ and ‘common’ are clearly and unambiguously defined. That is:

* Exactly what words fall into the category of “clear and common”
* Is there a list that I can choose from?
* Do all my words have to be clear and common- or just some?
* What about proper names (my name or the name of my city or state or county or the name of my school)? Some are common names, but most are not.
* How would I go about determining if my site conformed in a way that all testers would come to the same conclusion that I passed or failed.
* I cannot see how this could be tested – or that anyone could **objectively** and **reliably** determine that something had passed or failed – and I could not find any guidance on this in the supporting documentation.
* Presumably this accessibility guideline would/should meet its own requirements for simple language. (Or is this a technical document that would be subject to the exception?)

#### The exception for technical documents

This provision has an exception as follows:

**Exception**: Technical documents that provide a clear language summary

The exception would be normative – so also need to be clear and objective. This would first require a definition of a technical document. Is this a technical document? Clearly a paper on quantum physics would be. How about an advanced cookbook? Is the word “sauteed” a clear and common word? How about knitting? Can one have a site dedicated to it that does only uses plain language and no jargon? Or wood working? Or fishing? Or any site on any topic that is not intended for young audiences? Almost any topic is going to have jargon and words that are not understood by someone unfamiliar with that topic. Where is the line? How does one know if they have crossed it (to the good or bad side) (to the conforming or non-conforming side).

For technical documents a different rule is provided (labelled as an extension). But it is unclear what constitutes a summary. For the site on quantum or string theory – would the following summary qualify? “This is a site about what everything is made of.” For the WCAG 3.0 - would the following be sufficient summary “This is a document about how to make web pages easier to use by people with disabilities.”? That is clear and accurate but would not seem to be very complete for someone with a disability that pulled up the guidelines to find out what they contained. But that would meet the letter of the exception. And everything after that is opinion – and if we asked everyone on the committee what constituted an adequate summary, I would wager we would get as many different opinions as there are members – or more.

As I mentioned – this is really hard to do.

As an exercise – I suggest that each person in the working group look at their own website or that of their organization and apply this provision to the first 5 pages. Several of us tried this in WCAG 2 and it was a real eye opener. I thought it would be very hard. And I completely underestimated the task. And never did find a means to figure out what which words I could use. I even did a study of reading level measures and discovered all the limitations and ways they fail to address this.

I have no idea how to meet this provision. I would love to see the results of the application by the working group members to their sites. That is a good test for this provision – and for all new provisions. (In WCAG 2.0 we had a rule that we needed 3 implementations (sites) that met all of the guidelines before we could advance them. It was a lot of work, but it was highly educational and led to changes in the guidelines.)

#### Use of the technical provision to dodge this provision

Note that without a clear definition of “technical document” this provision could be easily met by simply putting a single sentence summary of each page, at the bottom of the page.

Can one say that anything that somebody can’t understand is too technical for them?

### 7.3 captions



The first line is again labeled as a guideline that reads "Provide captions and associated metadata for audio content".

* Because is non-normative, it does not need to be testable. But again I would suggest thinking about making these Out-come oriented – or dropping them and going straight to the Outcome. The imperative is confusing if it is not normative.
* The phrase “captions and associated metadata" should be a link so that someone can go find out what that is. Metadata is very clearly not a common word and very few people actually know what it means. (I am actually looking forward to a plain language explanation of metadata that is clear and easy to understand. I remember when I first encountered the word and tried to figure out what it meant many decades ago…)

The outcome then is "translate speech and nonspeech audio”

* First of all, this is not outcome-oriented unless the intent is that the page literally is a page that translates speech and nonspeech audio. That is, it literally is a page you talk to and it will translate speech, and also listen to all the audio in the background, and translate that as well.
* Since that is not the intent, I presume this means something like “provides a translation of speech and non-speech audio”. But that would be imperative and not outcome-oriented.
* Perhaps it meant to say, “A translation of speech and nonspeech audio is available".
	+ This would be an outcome and would be testable if they defined what translate means. To me translation usually refers to changing something into another language. The description below sort of explains what it means but still uses the word “translates”.
	+ If the text below is a rationale, it's not written as a rationale but rather it's written in a longer form of the outcome. These two should not be confused or conflated. Only the phrase above is marked-up as being the outcome. The phrase below is… I'm not sure what.
	+ Also – I note that the last text goes on to say "user agents in APIs support the display and control of captions".
		- It's not clear if that is a different requirement?
		- Is that a different or additional outcome that is expected - than what is specified in outcome?
		- I would note that these is not things that are under the control of the author (user agents and APIs). So, it can't be an outcome that the author is responsible for.
		- Is this simply a "and by the way" comment ? Or is it meant to be a separate outcome that should be listed as a separate outcome for a different audience than the author? If so, should it be a separate provision and tagged as applying only to user agent and API generators?

I would suggest something like the following

1. Consider dropping the guideline imperative
2. Change the provision to be outcome-oriented, something like
	1. “Speech and non-speech audio is available in a visual form, and in a form that is readable by assistive technologies.
3. Change the rationale to something like:
	1. Making the audio available in visual form makes it available to people who cannot hear or understand the audio for some reason. Making it available in a form that is readable by assistive technologies, allows others who cannot see or hear, or who cannot hear well enough or read well enough to have the content transformed into a form they can perceive and understand – including translating it into other languages or, in the future, into simpler language.
4. Move any requirement for user agents and API creators to a separate provision – that is marked as applying only to them.

I also note that the last paragraph says e.g. captions. This is fine if this is a rationale. If this was meant to be normative, then examples are not allowed. The statement needs to say exactly what is required. If an example is provided in the normative statement, then it is not clear if that is sufficient (e.g. nothing else needs to be done) or it is just one of a longer (unspecified therefore unknown to the author) list of things that need to be done. Do I also need to translate it into braille and symbols and other languages?

The outcome or objective part of this would similarly need to tell people exactly what formats are required or not required. Otherwise, it's untestable. The author is unable to determine if they pass or fail. And it would be impossible for others to judge reliably and consistently whether it passes or fails.

## Recommendations – possible solutions

1. **Rename the document and shift it to being a guidance document**
	1. the easiest way to preserve all of the value of this document is to rename it so that it is not trying to be a document in the WCAG line -- that is developed for use by people wanting objective standard that could be used in regulatory or other judgmental or criterion based evaluations ( e.g. as a requirements documents for purchasing or for use inside a company, inside an organization, inside a country, or internationally.)
	2. Such a document could be very valuable as a basis for continuous quality improvement for accessibility.
2. **Alternatively, one could decide to target it as an objective standard that could be used for in regulatory or other requirements use**
	1. But this would means having to follow the current WCAG model for selecting normative provisions (see below)
	2. This would allow the document to be used as a continuation in a WCAG line but would defeat many of the goals of the group putting this document together.
	3. The criteria for inclusion of provisions in WCAG 2.2 (from the working group wiki) are as follows:

Success Criteria shall:

1. Address a situation where a user with a disability will be disproportionately disadvantaged (as compared to a user without a disability) if the criterion is not met.
2. Be testable through an automated process or by a manual process conducted by an individual evaluator, and any tools required to test it are available prior to Candidate Recommendation stage.
3. Provide consistent results from different testers (e.g. 8/10 testers agree). This can be assessed through inherent logic or proven through examples.
4. Describe the specific condition required to meet the criteria, not the method to address the criteria.
5. Utilize the WCAG 2.x A/AA/AAA level structure.
6. Apply to all content across all websites unless preconditions for the application of the success criteria are explicitly identified (e.g. "except interruptions involving an emergency").
7. Apply across technologies to the greatest extent possible. (Technology-specific issues should usually be addressed in Techniques.)
8. Avoid creating a requirement for something that is already required by an existing Success Criterion.

Success Criteria should:

1. Be as broad as possible, but specific enough not to become a 'catch-all' for any given requirement.
2. Not require testing that is believed to require very large manual efforts. If a tool is very likely to be available soon after publication that makes the testing more efficient, this factor is less important.
3. Use glossary definitions to simplify and shorten all Success Criteria for shared or ambiguous terms.
4. **Since the guidelines are targeted to 3 different audiences (authors, user agent developers, and authoring tool developers) either:**
	1. **Break the provisions into 3 sections** – with each audience having a section they follow
	OR
	2. **Clearly separate provisions** meant for one audience from provisions meant for another **and tag or label them as such**
		1. So that each group and find what applies to them
		2. So that each group can meet the relevant provisions because they do not contain parts that are outside of their control
5. **Ensure that no provision is included until the working group goes and samples a good cross-section of the major sites, and the working group itself is able to rewrite them to show how they would/could conform. Implementations on at least 3 very different sites should be done for each and for all of the provisions (all the provisions applied to a meaningful number of pages representing the different page types for each of 3 sites).**
	1. This can prevent provisions that are not implementable from showing up in the guidelines
	2. It can help identify places where exceptions or additional conditions or allowances are needed (in WCAG 2 we learned much from this exercise.)

If the working group can't figure out how to redo the sites to comply – then it is likely that it is beyond the abilities of most web developers. And if it is too much work to do this for that limited number of pages – by the experts in the working group, then it is likely too hard to do.

1. **Continue with your Outcome-oriented approach.**
	1. Provisions should state what is desired – rather than what the content, tool, or agent author should do.
	2. How to do it should be in separate document that can be updated as technologies change. Not only how, but even ‘what’ needs to be done will change over time.
	3. This allows the provisions to remain stable – which is a key requirement for inclusion in regulations that may not change for 5 or 10 years sometimes.
2. **Use a consistent, straightforward and labelled format.**
	1. The three examples do not have exactly the same format. Also, they have sections that seem to be repetitive in functions sometimes.
	2. Suggested format {if you are going for a testable guidelines. If not then you can ignore the “normative/ informative” labels below – but this might still be a good format.)
		1. **Short name for provision**
			1. (makes it easy to refer to them in conversation and to find the one you are looking for
		2. **Provision**
			1. (In bold or highlighted or boxed text)
			2. This is the normative part of the provision
			3. Shortname can be in front of or above the provision
			4. The provision is in outcome format
			5. The provision does not refer to users or what users can or should be able to do – but rather what the should be true about the content, agent, or tool
		3. **A Rationale or benefit statement.**
			1. This is non-normative, informative
			2. It is below the provision and is labelled.
				1. To follow standard format for standards it probably should be labelled as NOTE:
			3. This note can use the word users and explains the purpose or benefit of the provision.
				1. This would help in understanding why the provision is needed.
				2. It can help readers to understand the different groups that can benefit from a provision (though that should be done in more detail in a separate document that discusses each provision in detail)
			4. Examples:
				1. NOTE: This allows users who are unable to perceive and / or understand the non-text content to determine its meaning.
				2. NOTE: Text alternatives to audio content allows users to access the information in the audio when they cannot hear the content or hear it clearly (e.g. if they are hard of hearing or deaf, or if it is too noisy for them to understand). Having the text alternative be accessible to assistive technologies allows it to also be translated into other languages, or into symbols, braille or other forms now or in the near future.
				3. NOTE: What are called captions in some countries are called subtitles in others – even if in the same language as the audio.
		4. **A link to further resources**
			1. This allows users who are unable to perceive and / or understand the non-text content to determine its meaning

Respectfully submitted

Gregg Vanderheiden

01-March-2021