

Accessibility and E-Resources: using a rubric to score platforms on their probability of inaccessibility

This document is adapted from a presentation given at ER&L 2022 by Hana Levay and Sidonie Devarenne.

Last edited: Jul 11, 2022

What is web accessibility?

The World Wide Web consortium (W3C) [defines web accessibility](#) as:

Websites, tools, and technologies that are designed and developed so that people with disabilities can use them. They can perceive, understand, navigate, and interact with the web, as well as contribute to the web.

How do we know if web content is accessible?

The W3C has developed [Web Content Accessibility Guidelines \(WCAG\)](#). These guidelines determine if web content is accessible when it meets certain success criteria. The current WCAG version is 2.1. WCAG version 2.2 is currently under development. A release date has not yet been set, but it will include several useful changes.

WCAG has 13 guidelines, and each of those guidelines has multiple success criteria. Each success criteria has an assigned level of A, AA, or AAA. Level A criteria are core criteria required to meet the minimum acceptable level of accessibility. The AAA level—the top level—are criteria that are important but can be more of a reach for many vendors. It's very unusual that a website would satisfy all of the AAA levels of success criteria. The usual consensus is to look for WCAG 2.1 AA levels of accessibility from a product.

Let's look at one of these guidelines in particular:

WCAG Guideline 1.2 is for time-based media, such as pre-recorded or live video. The goal is to “provide an alternative to time-based media.”

One of the level A success criteria for this guideline is “captions provided for pre recorded video.”

One of the level AA success criteria for this guideline is, “captions provided for live video.”

One of the level AAA success criteria for this guideline is “sign language interpretation provided for pre-recorded video.”

In this example, each success criteria level steps up to a more accessible feature. While it is fairly common to see that a vendor has provided captions for pre-recorded video, a level A criteria, we very rarely see that a vendor provides sign language interpretation, a level AAA success criteria.

Guideline 1.2 has nine total success criteria and each one is assigned a level of A, AA, or AAA.

Why do we need to know if web content is accessible?

Besides being an important civil rights issue, there are federal guidelines and often state guidelines that require federal agencies or state agencies to develop, procure, maintain, and use information technology that is accessible to all people. There have been many legal cases, including the recent example *Payan vs Los Angeles Community College District (LACCD)* that ruled that the LACCD must discontinue use of inaccessible library databases. By that, the judge specified that library databases must conform to WCAG 2.1 Level AA success criteria to be considered accessible. That's the middle level: from the previous example, that would be "captions available for live video." This ruling gave many libraries a bit of a panic. While libraries had also sought to provide accessible resources to users, they now had a specific set of standards to meet in order to call a resource accessible.

Which library resources are not accessible?

To discontinue use of inaccessible library databases, we first need to know which databases are not accessible. There are several ways to approach this problem. We can do manual testing. For example, there are a variety of accessibility tools that are freely available that you can use to manually test web content, including browser tools and plugins like [Accessibility Bookmarklets](#) and [Tota11y](#). These tools highlight problem areas on web pages such as inadequate color contrast; inconsistent heading structure; missing alt text; etc.

There is testing with screen readers, where you navigate the web page using screen reader technology, to find issues with page structure or problems. You can use keyboard navigation testing, where you navigate the web page using only the keyboard controls, and try to successfully accomplish the various tasks a user would be attempting on the resource.

Another option is to assess [Voluntary Product Accessibility Templates \(VPATs\)](#).

A VPAT is a vendor-provided document that summarizes the accessibility features of their product. There are a few different versions of the VPAT, but I will be focusing on the [VPAT 2.4 Rev WCAG](#) version which is the most commonly used version for third-party library resources. This version itemizes each WCAG success criteria for each of the 13 guidelines. The vendor is responsible for declaring if their product supports each of those guidelines, and at what level, but there is no mechanism to ensure their claims are accurate or any repercussions if statements turn out to be inaccurate.

We can collect these VPAT reports and we can manually assess each resource for each success criteria, but all of these options are huge projects. Most of us don't have the expertise, funding, or time to get started. Additionally, it can be overwhelming to someone new to accessibility.

This is where many of us are now.

How do we get started?

Let's remember our goal: we want to improve the accessibility of subscribed third-party library resources. We have these constraints:

- VPATs are complicated, hard to read, and many are actually unreliable
- Libraries, of course, have no control over the development of the third party resources; we can't actually make any fixes.
- Libraries subscribe to many resources, so it's a very large project.
- Accessibility testing tools are often difficult to learn.
- Finally, many libraries lack funding and time to address these issues. We often have to do this work as part of our regular workload with no additional resources.

At the University of Washington, we decided the best balance between our limited resources and our need to assess accessibility would be to use a combination of keyboard testing and VPAT assessment to find which electronic resources are most likely to not be accessible. Then, we negotiate with the vendors of the likely inaccessible products and work with them to encourage them to improve the accessibility of their products. While this approach does not thoroughly test products for accessibility, it does give a good idea of which vendor products we should focus on.

Keyboard testing

We chose to use keyboard testing to manually test our subscribed resources because it is easy to learn, easy to use, and there is no software to download. Generally speaking, if a website is not accessible using keyboard navigation, it will also fail to work with various assistive technologies.

The W3C says "an accessible website does not rely on the mouse. It makes all functionality available from a keyboard. Then, people with disabilities can use assistive technologies that mimic the keyboard."

The basic commands for keyboard navigation are using tab to navigate forward through the page or shift-tab to go backwards through the page, using enter and spacebar to activate buttons or links, and arrow keys to browse menus. There's not much more to it. It's very easy to learn and easy to use.

Keyboard testing cheat sheet



- Navigate web pages using only the keyboard
- [#nomouse challenge](#)
- **Tab** – move to the next link, form element or button.
- **Shift+Tab** – move to the previous link, form element, or button.
- **Enter** – activate the current link or button.
- **Space** – check or uncheck a checkbox form element. Will also activate a button that currently has focus.
- **Up/Down arrow keys** – move between radio buttons or, in some cases, menu links.
- **Right/Left arrow keys** – in some cases, move between menu links or adjust sliders in audio and video plugins.
- **Escape** – Close the current modal dialog or dropdown menu and return focus to the element that spawned it.

Once we settled on keyboard testing, our goal was to use only the keyboard to see if we could accomplish the main actions a researcher would d.:

- Search,
- Select a search result,
- Interact with a search result, like turn pages or scroll,
- For streaming media, play and pause the content.

Keep in mind we did not test the content itself at this time, just the interface. Once you become familiar with this method of testing, resources can be tested in about ten minutes per resource, so this is an efficient process.

This test only finds if common actions can be accomplished. It does not test for accessibility beyond that. But given our constraints, keyboard testing gives us a pretty good picture of the functional accessibility for the resource.

At the last count of the University of Washington keyboard test results, 90% of the 606 resources tested passed the keyboard navigation test. So, we know that we need to look harder at those 10% of resources.

We published the results of our keyboard testing on a [public web page](#). This testing process is ongoing. We continue to test and retest resources and update the results. One could decide to end assessment here, but at the University of Washington, we decided to incorporate VPAT analysis as well.

What about VPATs?

The challenge with VPATs is they can be long, complicated, and overwhelming to read, especially for anyone new to accessibility. Also, they may be unreliable: just because a vendor provides a VPAT does not necessarily mean that their resource is accessible. Ideally, a third party accessibility consultant will be the one filling out a VPAT for a company, and testing for each success criteria. But in reality it is often filled out by a vendor's marketing team who have very little accessibility expertise—they've just been asked to fill out the form.

Because VPATs can be so long and complicated, we have decided to look at just a few key details in each VPAT. These details were drawn from a UW Accessible Technology webinar [Accessibility in Procurement](#), which suggested that these are the most important elements of a VPAT and give us a very good sense of the overall quality of the VPAT. This is assuming we are examining a current VPAT 2.4 Rev WCAG.

These key elements are:

- Metadata: name, date, contact, applicable standards, and evaluation methods used,
- Section 1.3.1: "information and relationships,"
- Section 2.1.1 "keyboard testing," and
- Section 4.1.2: "name, role, and value."

The quality of how these elements have been completed demonstrate a lot about the quality of the VPAT itself.

Some VPATs are publicly available. In the best case scenario, vendors make their VPATs available on their site. Some resources have a link to their accessibility policies and documentation in the site header or footer. Others link to this information in the About, Help, or Knowledge Base sections of their site. VPATs for platforms owned by large vendors may be located on the parent company's site.

If it's not clear how to find the VPAT on the site, try a site search or try using Google to locate it. You may also find that another library or library consortium has published the vendor VPAT on their site, though you'll need to make sure it is the most recent version.

Finally, some vendors make their VPATs available by request only.

To assess the collected VPATs, we extracted the key elements mentioned above. We directly copied the text for the selected VPAT sections into a spreadsheet, with the exception of Evaluation Methods. We tried to summarize evaluation methods information as:

- "Blank" if they left it blank,
- "Minimal" if they just said something vague, often with phrasing such as "tested with appropriate technologies"
- "Full" if they itemized each tool they used to test.

We had some challenges collecting this data. We found that many VPATs are not actually filled out completely. Also, sometimes it was hard to use the scale to summarize the evaluation methods used because different vendors provide different levels of detail when describing the evaluation methods. They might just say they "used assistive technologies," versus naming specific products, or that testing is "based on general product knowledge," which is very vague. Ultimately, we found a range of detail provided in VPATs. Some indicated a thorough assessment of their product, while others really didn't tell us much.

Once we have the VPAT metadata collected, how do we assess it? We took the results of our keyboard testing and our VPAT assessment and put them together into a rubric with four elements. Each element was scored, where zero is the best score for each section, then points are added where each point is a mark against the resource.

The first element is the "VPAT completion score," or "how does it look?" This considers how the VPAT looks without considering the content: Is it out of date? Do most of the comment fields have text? For example, one time we just received the blank VPAT template back from a vendor with no edits at all. In that case, the VPAT would receive a score of 3 on this section. Ideally, VPAT should be current and have all of the desired metadata elements filled out and with plenty of text in the comments sections. This type of VPAT would be scored a 0 for this section.

The second element is the "VPAT success score", or, "what does it say?" In this element, we score based on the actual text of the VPAT. A VPAT that includes mostly "supports," good descriptive text in the comments fields, and a fully itemized evaluation methods section would be a 0. Conversely, if it includes mostly "does not support," the comments fields are empty, and no evaluation methods are listed, then this section would be rated 3.

The first two elements also allow for a result on a scale from 0 to 3. It is not always clear how to score on the scale, so some experience and judgment is required.

VPAT + Keyboard Testing Rubric

VPAT Completion score: "how does it look?"

- 0 - Complete VPAT
- 1 - Many edits but no comments
- 2 - Some edits but not helpful
- 3 - None/out of date/blank/default template

VPAT success score: "what does it say?"

- 0 - Full evaluation method, most "supports" with comments
- 1 - Some "partially supports" with comments
- 2 - All "partially supports", or, no comments
- 3 - Mostly "does not support", no comments, ineffective evaluation methods

VPAT 2.1.1 "Keyboard Testing" and keyboard test agreement:

- 0 if they agree
- 3 if they don't agree

Keyboard test:

- 0 if passed
- 3 if failed

Overall score: higher score correlates to higher risk of inaccessibility (range of 0 - 12)

For the third element of the rubric, we bring in the keyboard navigation test result. The resource gets 3 points if it fails the keyboard navigation test and 0 if it passes.

For the fourth and final element of the rubric, we check to see if VPAT section 2.1.1 "keyboard testing" success criteria and our keyboard navigation test agree with each other. If the VPAT shows "supports" on this item but failed our keyboard test, this element would be scored 3 points.

For the overall result, on a range of 0-12, the higher the score correlates to the higher risk of inaccessibility. Generally speaking, a score of 6 or higher would indicate a resource to be looked at more closely.

VPAT Examples

If vendors leave the Remarks and Explanations field blank, like this example from the American Psychiatric Association VPAT, or simply paste in the official success criteria, we can't really evaluate how closely they assessed their resource:

Criteria	Conformance Level	Remarks and Explanation
1.1.1 Non-text Content (Level A)	Supports	
1.2.1 Audio-only and Video-only (Prerecorded) (Level A)	Supports	
1.2.2 Captions (Prerecorded) (Level A)	Supports	
1.2.3 Audio Description or Media Alternative (Prerecorded) (Level A)	Supports	
1.3.1 Info and Relationships (Level A)	Supports	
1.3.2 Meaningful Sequence (Level A)	Supports	
1.3.3 Sensory Characteristics (Level A)	Supports	
1.4.1 Use of Color (Level A)	Supports	
1.4.2 Audio Control (Level A)	Supports	
2.1.1 Keyboard (Level A)	Supports	
2.1.2 No Keyboard Trap (Level A)	Supports	
2.1.4 Character Key Shortcuts (Level A 2.1 only)	Supports	
2.2.1 Timing Adjustable (Level A)	Supports	
2.2.2 Pause, Stop, Hide (Level A)	Supports	
2.3.1 Three Flashes or Below Threshold (Level A)	Supports	
2.4.1 Bypass Blocks (Level A)	Supports	
2.4.2 Page Titled (Level A)	Supports	
2.4.3 Focus Order (Level A)	Supports	
2.4.4 Link Purpose (In Context) (Level A)	Supports	
2.5.1 Pointer Gestures (Level A 2.1 only)	Supports	
2.5.2 Pointer Cancellation (Level A 2.1 only)	Supports	

Similarly, a VPAT doesn't tell us much if we don't know any details about their evaluation methods, like these examples from GeoScienceWorld and Kanopy:

Evaluation Methods Used: Automated Testing; Exploratory Testing

Evaluation Methods Used: General product knowledge

The VPATs we found helpful gave us information that showed the resource had been thoroughly tested. In these, the Remarks and Explanation sections were filled out with specific detail about the resource, as in this excerpt from the Wiley VPAT for success criteria 4.1.2 “name, role, and value”:

<p><u>4.1.2 Name, Role, Value</u> (Level A)</p>	<p>Partially supports</p>	<p>For most user interface components, the name and role can be programmatically determined; states, properties, and values that can be set by the user can be programmatically set; and notification of changes to these items is available to user agents, including assistive technologies. However, there are minor exceptions. These include:</p> <ul style="list-style-type: none"> • Tooltip controls are not exposed to assistive technologies. • Modal dialogs lack appropriate roles and ARIA attributes. • Autocomplete functionality lacks appropriate roles and ARIA attributes. • Custom <select> widgets lack appropriate roles and ARIA attributes. • Error states of form elements is not made programmatically available to assistive technology users. • Tabbed navigation has not been assigned the correct roles and ARIA attributes.
--	---------------------------	--

This vendor has stated that it partially supports this success criterion and then includes many details in the remarks and explanation field about exactly what it supports and any exceptions. It is clear that they thoroughly tested it.

One thing to notice here is that we almost don't even need to read all of the text; the fact that they have so many comments supports the idea that they have put some time into testing and

filling out the VPAT.

Similarly, in the “evaluation methods used” section when a vendor shares specifics about their testing it helps us to know that thorough testing was conducted, as in this excerpt from the Alexander Street VPAT:

Evaluation methods used:

Alexander Street code is checked for accessibility using a range of visual, automated, and manual checks. For example, colors used throughout the interface have been tested against the World Wide Web Consortium's (W3C) standards for luminosity contrast, color difference, and brightness difference. Other tools used to test accessibility include:

- NVDA with Firefox
- WAVE (web accessibility evaluation tool) provided by WebAIM
- Color Contrast Analyzer provided by WebAIM
- Manual keyboard checks

Thoroughly tested resources might even be more likely to call out their product as being in partial conformance rather than fully supports because they’ve really taken the time to investigate and report on results

You can see a range of scores from the rubric as determined at the University of Washington. When the “overall score” column is sorted from lowest to highest, you can see a list of resources that are most likely accessible:

VPAT Platform/Vendor	VPAT Completion score: 3 for out of date/blank/default template; 2 for some edits but not helpful; 1 for many edits but no comments; 0 for a complete VPAT	VPAT success score: full evaluation method, all "supports" with comments - 0; some partially supports with comments - 1; all partially supports or no comments - 2; does not support, no comments, ineffective evaluation methods - 3	VPAT and keyboard test agreement: 0 if they agree, 1 if they don't agree	Keyboard test: 3 if failed or untested, 0 if passed	Overall score: higher score is higher risk (range of 0 - 10)
Alexander Street	0	0	0	0	0
IBIS World	0	0	0	0	0
JSTOR	0	0	0	0	0
ScienceDirect (Elsevier)	0	0	0	0	0
Social Explorer	0	0	0	0	0
Adam Matthew	0	1	0	0	1
Brill	0	1	0	0	1

Again, keep in mind that we are not testing the resource to see if it is fully accessible. The goal here is to get a sense of which resources are probably not accessible. The result of the rubric is a ranked list in order of inaccessibility. This allows us to spend more time on vendors with the products that are most likely inaccessible.

Here is the same list with “overall score” sorted from highest to lowest, highlighting the resources that most likely have accessibility issues:

	Rubric					
VPAT Platform/Vendor	VPAT Completion score: 3 for out of date/blank/default template; 2 for some edits but not helpful; 1 for many edits but no comments; 0 for a complete VPAT	VPAT success score: full evaluation method, all "supports" with comments - 0; some partially supports with comments - 1; all partially supports or no comments - 2; does not support, no comments, ineffective evaluation methods - 3	VPAT and keyboard test agreement: 0 if they agree, 1 if they don't agree	Keyboard test: 3 if failed or untested, 0 if passed	Overall score: higher score is higher risk (range of 0 - 10)	
WOIS	3	2	0	3	8	
ASTM Standards & Engineering Digital Libra	3	3	1	0	7	
Associated Press	3	3	1	0	7	
HSUS - Cambridge UP	3	3	1	0	7	
New Play Exchange	3	3	1	0	7	
PrivCo	3	3	1	0	7	

Conclusion

At the University of Washington, we have completed the full rubric assessment on 53 resources at the time of this writing. We found about 15% had a high score, which means they are more likely to be inaccessible.

These results are more nuanced than just a pass/fail. When we combine the VPAT metadata and the keyboard test, we create a ranked list of resources and vendors to find which vendors are more likely to have inaccessible products. This helps us refine our scope, so that moving forward, we can focus on communicating with the vendors who scored the highest.

We can work with these vendors. We can ask them to add accessibility language to the license. We can ask for a one-year license with the intent to renew next year only if accessibility improvements are made in that time. We can include accessibility findings as a factor when considering products to license or to cancel.

Overall, we can inform vendors that accessibility impacts our decision-making. Vendors need to work with us to improve accessibility. They are our partners in this work, not our adversaries. In some cases, we have found that a vendor is not aware that their product is not accessible and they appreciate working with us and learning how to improve their products. In another case, our decision to cancel a resource because of its inaccessibility was motivation for them to improve their product.

Additional resources

The [Library Accessibility Alliance](#) is an alliance of about 100 libraries who work together to

promote equitable access to library resources. The LAA provides [model license language](#) to use in new licenses and a [toolkit](#) for getting started. They also do thorough, high-level accessibility tests with third-party accessibility consultants. These reports are available on their [testing page](#), which also shows if a vendor has responded to the reports.

The LAA also provides an [e-resource evaluation suggestion form](#) to suggest new resources to test.

We also recommend:

- [WCAG Guidelines](#)
- [UW keyboard testing results](#)
- [VPAT + Keyboard Testing Rubric Template](#)
- VPAT Repositories
 - [CUNY](#)
 - [Shoreline Community College](#)
 - [RAILS Libraries](#)
- Webinar: [Accessibility in Procurement](#) and [slides](#)