

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

Present

- James Craig
- Matt King
- Michael Fairchild
- Simon Pieters
- Isaac Durazo
- Rick Waldron

Slides

<http://a11y.pro/about-aria-at/>

Action items and insights

- Tests
 - Unlike other SRs (JAWS/NVDA), VO doesn't have modes, but different navigation methods. All/most navigation methods are available at all times. I think what Matt was getting at here was whether the SR is in a state where it "intercepts" certain keypresses. For example, Windows screen readers intercept Tab, Arrow Keys, and most single-letter keypresses when "browse" mode is on. VoiceOver does not intercept any standard keypresses with default settings.
 - There can be user settings affecting results.
 - We assume default settings.
 - ARIA groups surrounding checkboxes are currently required (by non-normative APG, not the normative ARIA spec?) to behave like fieldsets. James thinks this is problematic and could result in problems with verbosity. It makes sense to convey group boundaries, but not necessarily announcing all ancestor group names when navigating between elements, for example.

Bocoup

PO Box 961436
Boston, MA 02196
617 379 2752

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

- James suggests starting with simple tests first. For example, a normal checkbox without a group.
- It would be interesting to test native HTML widgets.
 - This is currently out of scope, but we could increase the scope (modulo funding).
 - Difference between native widgets and custom widgets
 - With a native checkbox, you can swipe up/down to change its state. Not with a custom checkbox.
 - Slider
- Presenting test results
 - Watch out for penalizing for things that aren't supported (and there isn't consensus on whether it should be supported).
 - Consider not counting results when there's no platform equivalent.
 - (Should we not record that as a bug for the platform?)
 - Overall scores for things where everyone agrees should happen seems OK/good.
- Process
 - In case of disagreement, it could result in changes to:
 - ARIA spec
 - APG
 - AT implementation
 - The test (remove it, change it, change priority (must have, should have, optional)).
- Project timeline
 - Consider conflicts with conferences:
 - Global Accessibility Awareness Day
 - Google I/O
 - Apple WWDC
 - Microsoft EdgeConf
- Stakeholder support
 - James thinks he can review test plans twice annually.
 - James wants to review the initial set of results, but not every single change.

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

- Keep James informed about the next stage in the process.
- Feedback on the presentation
 - Scope slide: expand to native HTML semantics. (Done.)
- Feedback for APG TF
 - Consider renaming ARIA Authoring Practices to Accessibility Authoring Practices.
- Feedback for the ARIA-AT CG
 - Use living documents. It should be easy to correct mistakes.
- Things that should be clarified
 - Relationship with ATTA (by Jonie, Shane).
 - ATTA is an automatic test framework for OS-level accessibility API.
 - ARIA-AT is for AT behavior for manual tests.
 - Why desktop screen readers are being tested first. Why not mobile.
 - Problems are richer in desktop. Also no clear mobile scope within the ARIA WG. We only test on desktops and don't know what to do about that.

Raw log

MK: Reviewing outcomes of the meeting

MK: Reviewing Vision

J: I have some questions about how this would change how SRs work, but will save questions until the end

MK: Ok. Not necessarily that SR will change, but let's talk about it later.

MK: So, the high level plan is to define expectations for AT in the form of test plans and then measure the support for those expectations.

MK: Reviewing principles for the project

- Vendor neutral

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

- Open
- Consensus

MK: Reviewing Scope: focus on desktop SRs for now but still plan to go beyond this in the future and also branch out on other types of AT technologies and potentially developing some public APIs for automation that could be integrated into AT technologies

J: Why start with desktop browsers instead of mobile browsers? Is it because of more inconsistencies in the desktop?

MK: We can't actually tell you if there's more inconsistencies because there's no data. The problems are richer in desktop. Also we don't have a clear mobile scope within the ARIA working group. We only test on desktops and don't know what to do about that.

J: Yeah. Complicated answer. I think I agree with the scope of focusing on desktop SRs for now.

MK: there's little lack of vendor neutrality in the definition of scope itself because we didn't include Microsoft SR. That was a bandwidth decision. I'm hoping that's not a problem

J: Have you talked to Rossen about it?

MK: No, but we will be presenting to them. I sent the inquiry initially to Bogdan but haven't gotten back to me. He's the Aria working group rep. If you think we should talk to Rossen maybe we should.

J: I'd probably CC both Rossen and Bogdan.

MK: Ok. Progress to date will be covered by Simon

S: So the AriaAT community group was formed in 2018. We, Bocoup, have been contracted by FB to research and prototype our first version of a test runner and see how it could work.

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

S: This year we are spending time building a production ready testing system and the idea is to have tests for everything in APG

S: Right now we have 3 tests for 3 design patterns in APG and still in the research phase for creating test runner and ambition for authoring tests

S: We have a project plan. The important dates are:

1. **March 20:** have complete plan for how stakeholders will participate and how we plan to write tests
2. **May 22** intent to have first draft of all tests, which is important because we want to get feedback
3. **June 12:** we'll run a pilot test of the tests
4. **June 26:** we want to have a complete run of all tests with 6 screen readers combination in scope

J: Feedback on "Progress to date" slide. One of the notes to share was that John Diggs and Shane McCarron worked on something like this where they had a single form of test but multiple tests runners per platform (?). Some of them wrote a platform-side accessibility API crawler that acted like a server and would talk to the test runner and the runner would activate the JS in page to change things... so just be aware some of this still might exist.

MK: John took the ATTA and we have one for MS8, for Orchha and... it looks like the accessibility API content from OS-side... so that's all that is.

MK: It's good for testing what ARIA requires, but it does not cover our goals, which is to test what happens from that point up through to the user experience.

J: Is this intended to be manual testing?

MK: Yes this is manual for now

J: Ok, thanks. I have a comment on the next slide (the one about dates and project plan)

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

J: Dates are potentially problematic. May 22-June 12 because May 21 is Global Web Accessibility Awareness Day and some major vendors have big conferences: WWDC, I/O, and EdgeConf. It might be a tight time to ask for feedback.

MK: Ok, hopefully resolved when we know who's doing what

S: we'll follow up to know when it is best to get feedback. We can reach out at different people depending on their schedule

S: Will share a link to dev plan later.

J: sound great.

S: Now sharing test results ready for checkbox in Voiceover in Safari...

MK: as a side note... let's take 3 patterns with ARIA patterns in it.

MK: Simon it might be good to share the example before the results

MK: didn't anticipate the checkbox example would be this difficult

MK: ok...back to the report

S: So the summary of tests. We have a bunch of different cases. Each case has one interaction and a set of expectations for interactions. This have different priorities divided by "must have", "should have" and "nice to have". For this group of test "must haves" is only expected and overall results is 84% in VO

MK: I want to point out one that there's some test related to leaving a checkbox group and that right now JAWS and VO don't say anything when a user leaves a checkbox group and goes to another element that's not in the group.

S: There are only 7 cases in this report's result

MK: Something happen then because there should be 16

S: I know JAWS results has 17 or 16

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

MK: Ok maybe that's the one. Referred to the questions James had early before

MK: the point is that there's certain behaviors that VO doesn't do and we are not totally comfortable with the "Nice to have" label because if there's something some SRs that others don't ...MISSING PART... Question is how a user is supposed to know that? How do we convey grouping info? What should the priority of that be? A should have? Or a nice thing to have?

J: I would recommend starting with a simpler example like a single checkbox. There's still gonna be differences based on how you navigate.

MK: We think we have those differences covered.

J: if we were to change something about VO that could change a lot of the expectations. You may want to consider making the updates to this. Somehow very flexible. The other thing about this example that complicates things is what you call interaction mode

MK: Let me cover both of those points. We came up with generic terms.

MK...In voice over we only have interaction mode

MK: One question for you is if you think we should test things with quick nav on and call that Reading more? Right now our current test for VO is all interaction mode tests... We are not testing reading mode because reading mode equivalents.

J: It's possible that's just a vocab mismatch. VO doesn't have modes. Anything's available at all times. We will refer to that as different kinds of navigation. Then there is VO navigation, which allows you to walk the AX tree, ~versus Tab navigation which relies on the system KB focus Both types of navigation are available at all times, but the user may have different expectations about what they hear based on the type of navigation they used to get to where they are.

J: We may speak an extra label...

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

MK: Right. It's a vocab thing and VO has quick nav on and off and when it's on still has other voiceover features.. So it doesn't affect VO operation.

MK: to get back at the fundamental point. About differences in behaviors... the way we wrote assertion was... we say a certain thing has been “conveyed”

MK: This is what we are going to rely on having expert testers. And of course we are gonna peer-review results but want to make sure that If conventional way of conveying something for a particular SR is for example “make a sound” or a word we regard that as the way of conveying that information and that is incorporated in how the assertion is interpreted.

MK: So we are not baking in to the assertions that things have to be done in a certain way.

J: This makes sense.

MK: So we plan to use defaults. Our thinking is test with defaults

J: Makes sense too. Part of the reason why I'm against this example is because I'm not sure I'll agree with some of the assumptions.

J: I would encourage to focus on simpler examples and work towards more complex

MK: Yeah we are starting with what we have in the APG.

MK: if there's an assertion with no general agreement one option is to remove the assertion and discuss within ARIA WG what our intent is for that role or state or property is.

J: that makes sense too.

MK: another is that we change the priority of it.

MK: Our goal for the “must haves” is that when a web author looks at the APG and look results for particular pattern and when they see that the “must haves” are not yet consistently supported in an AT they might want to question if they should rely

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

in that pattern and our goal is to get it to where all the AT support the “must haves”. This is where getting consensus is very critical

J: one more question regarding failing in that screen. Is that assuming all of them are failing or one of them?

MK: Yeah, the failing assertion is a list of all of the ones that did fail for that test. There will be passing and failing assertions if they are mixed

J: OK. If the long term intention is to store this in a DB format it;s better to get the results as granular as possible.

MK: This is built from a JSON file.

J: Ok

S: one aspect that’s changed from this test result and the newer test results we have for JAWS is the granularity of tests itself

MK: James you might be interested... we are using the web platform tests format for the tests themselves. Same stuff CSS working group is built on

J: great!

S: moving on to Goals for Reporting Test results

S: based on early concerns with wpt.fyi, we don’t want to have overall scores of tests results. Goal is for web devs to improve the quality of ARIA support and increase reliability and interop for how SRs work and for them to understand what patterns are well supported

J: I’m curious. It sounds like somebody said don't want to compare scores. Did someone make an objection?

S: no objections but based on experience with html5test that we know we don’t want to do that style because it brings wrong incentives. We want to focus in

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

celebration improving and collaboration and making sure we move to a more reliable platform for devs

MK: I have no problem saying we compile VO results into a single report for Apple but our goal isn't to publish like: "here's everything from VO", but if you are a web dev you go to APG and open a pattern and we'll show VO support, JAWS support, NVDA support in different browsers and show how well they do in for that pattern.

J: A couple of things I want to say is when Steve worked on html5 accessibility one of the things that VO was penalized for was things that webkit had not adopted.

J: Another thing to watch out for is "a percentage of things that everybody agrees should happen and are not" I think is a good thing to have in general. I think AT shaming is a good motivator.

J: but make sure you don't penalize for things that people don't agree on

MK: I think I want to get to where everyone agrees on the "must haves" and go through a backlog for the ARIA WG or the practices WK. There's some upstream dependency. We should not measure anybody on anything that has an upstream dependency

J: And this example too is one of those things. This is an ARIA group but by assumption you decided that it needs to behave like a Fieldset in this context. If that's the case that's news to me.

MK: Yeah, that is what's implied in the ARIA spec and what we are telling people in the practices, so if people don't agree then we have to revisit the spec.

J: group specifically around a set of checkboxes only or any checkbox exists inside any group all together. if there's a group that just happens to be a group and there's a bunch of content in it and it happens to be a random checkbox. being a little granular but i think you get the idea..

MK: Yes I get the idea. We need to dig into that

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

J: a Fieldset is a more limited type of group and we have that behavior and that behavior is generally agreed upon and we need to convey to us because of accidental containment

MK: Simon can you use a fieldset on ARIA checkboxes that ...MISSING PART...

S: Well fieldset is an element. We need to use that element

J: none native form elements?

S: Yes there's no requirement for what goes into a Fieldset so you can put anything

MK: I wonder if we should revisit the practice guidelines based on what you are saying James. There might be a reason to discourage people from using "group" in this way. Isaac please capture that thought. Don't want to forget this because one of the things that will happen as a result of this is... practices has been working on improving ...MISSING PART... we should really change what's in practices

J: Yah, so one of the things I want to say is that it might make sense to convey a group label if you cross that group boundary.

MK: VO command ...MISSING PART...

J: yeah I'll have to ask some people here...MISSING PART...

MK: dont know of command in VO to provide context

J: I'm Asking a rhetorical question. What's correct behavior? If what you describe is the correct behavior then should it be that reads every meaningful ancestor?. Not saying it should but where's the cut off

MK: right... in terms of working together going forward. Obviously a lot to discuss and things to work out. To what extent could apple be involved and help us get it right?

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

J: I see a bullet saying test and report twice annually. I can agree with that, I'm ok with that expectation. Weather half a week or something like that to jam through all of the docs. Seems reasonable to me.

MK: reviewing test plans before executing. Is that something someone from the VO team?

J: it will be me reviewing those. I can do that as part of my contribution to this effort. Again I would say start small. See how many questions I asked just about this example because of the oddness of the grouped checkbox example in the APG. It might be worthwhile adding a couple simple examples to APG, such as an ungrouped checkbox, and start with those.

MK: Yeah so the next one we have is combobox and Menubar, which includes a whole menu related stuff

J: right

MK: it is possible you can test a lot of these things in isolation but won't necessarily help people know if they can use this pattern.

J: it might be good to compare this to native elements. One of these long term challenges we are still working on inside the AOM incubation is the fact you can't swipe up and down on ARIA checkboxes. we think we are getting closer to agreement there. People don't realize this is still a problem and if they do they realize it is for iOS swipe up and down on a checkbox. The VO Nav or Quick nav equivalent to that also doesn't work on Mac because it's not a real slider. You could use these support docs to call that out and recommend native range sliders.

MK: yah that's a direction we want to take. Right now we want to measure support for ARIA and not necessarily for native semantics. I would like to include native semantics but APG doesn't right now. We could increase the scope of this project to do that...

J: The goal I want I was trying to convey here is that there's a lot of qweb authors that their initial decision point. They could've chosen a native element that has

Bocoup
PO Box 961436
Boston, MA 02196
617 379 2752

Meeting Notes: ARIA-AT intro for Apple



2020-03-02

better support and get more of it for free. And ARIA says that at the beginning of the spec.

MK: Yeah until I can get the TF....MISSING PART... authoring practices guide. I have that in the roadmap. WAI did assign someone to focus in that space. So for now getting this off the ground... the scope is explicit ARIA semantics and SRs because of the level of problem but I intend for this to be more encompassing and I think we fail to mention that on the scope side. Expanding to other AT but other HTML semantics and other ways to implement the same semantics.

J: like idea of changing APG to Accessibility Authoring Practices Guide

MK: Yeah I would love for every single pattern is the most native one you can possibly make and say **PREFERRED**. We are close to covering ARIA and initiated those conversations

J: Great! I want to see the first set of results. What would recommend is to keep this work close to a living doc as possible. The Accessibility API mapping spec is one of the ones for some reason Rec Track instead of Living Standard. We have requested to make it a Living Standard.

MK: Ok I think we've covered everything

S: Yeah

MK: we'll keep you informed of dates and milestones

J: keep me informed on the next stage of the project and keep the May 22 - June 12 dates flexible because it will be a busy time for many. May 21 is Global Accessibility Awareness Day, followed soon after by Apple WWDC, Google I/O, and Microsoft EdgeConf.