# Webinar: DAISY- Publishing, accessibility, W3C standards – where are we and how did we get here?

The recording and further details from this webinar are available at:

<https://daisy.org/news-events/articles/publishing-accessibility-w3c-standards-w/>

Date: 5/6/20

This is a Captioned transcript provided by CIDI to facilitate communication accessibility and is not a verbatim record of the session.

>> Richard: Hello everyone. My name is Richard Orme from the DAISY Consortium and I am your host for today’s webinar. We'd like to extend a very warm welcome to this session entitled Publishing, accessibility, W3C standards- Where are we and how did we get here?

This is the first of two webinars examining the revolution of born accessible digital publications. Part two will consider "The future of accessible publishing and standards “where are we going?. I'll remind you about that session again at the end.

But let’s get started! I'm thrilled to hand over to our panelists, who will introduce themselves and ponder "Where are we, and how did we get here?

>> Bill: Hi, I'm Bill Kasdorf. I'm a publishing independent consultant. I describe my consultancy as publishing infrastructure, editorial workflows, best practices and standards alignment and accessibility. But actually accessibility is a part of all of those things. So I shouldn't list it as a separate thing. That's what I do.

>> Luc: Hello, I'm Luc Audrain. I'm rehired from Hachette Liver. I have been working for many years. I'm happy to let you know what we did inside the W3C where I was the cochair of the publishing business group for some years, since 2017, since its creation. It's very a big honor to present the situation of where we are now.

>> George: My name is George Kerscher. I'm chief innovation officer with the DAISY consortium. And senior officer of global literacy with benetech. I also mention that I'm blind and use a screen reader. All of this access technology that is evolving in publishing is just wonderful to see.

>> Bill: Okay. This is bill. I'm first up today. So I will be talking about publishing technology. I will then hand it over to Luc Audrain who will give a description of the EPUB ecosystem that he has had a hand in building. Following Luc will be George and talking about baking in accessibility and then we will leave plenty of room for discussion at the end. So let's get going.

I've been saying for a few years now that accessibility is more accessible than ever. The publications are more accessible, the systems are more accessible and the devices that we use to access those publications are more accessible. Why is that? Because of how publishing technology has evolved. Step back a few years -- 50 years is more than a few to most of you, not to me. When publications were print bide letter press. Then they were consumed on the paper. Systems in producing them the writing was on a typewriter, manuscripts were mailed. The editors use red pencils to mark in the margins. Went to a type setting. The led was put in a printing press and books were printed and bound and shipped to bookstores and people went to bookstores and bought them. Are you kidding me? Is that how it was? Yeah. It's inherently inaccessible. Every aspect of that is inaccessible. You could read the book at the end of the day if you can handle it properly and turn the pages properly and see the print.

So you don't think I'm making that up. This is a picture of me in 1971. I was printing poetry chap books.

In the early digital era, typesetting did become digital. The manuscripts that were being type set were still on typewriters. Publishers used proprietary markup that was presentational. The markup said what content object looked like. It didn't say what the object was. When those marked up manuscripts went through a desiring and typesetting, the designers made up the codes for an individual book one by one. If they were using in design that came along later, making up the style names one by one. The typesetting did the name thing. So there was no consistency even though we now have digital files. It still resulted in a print book. So what that meant was Assistive Technology just had to start from scratch with the print book to make an accessible product. They used very specialized systems that making the accessible product use specialized systems and coding and consuming the accessible product also use specialized systems and coding.

So fast forward to today. This is why I'm just so enthusiastic about where we have gotten and echoing George's comment about how publishing technology has advanced and is inherently made accessible.

Every step of the workflow is digital. Even producing print is a digital workflow. I did a book called the Columbia guide to [inaudible]. It was everything about publishing. Everything about that book was digital. Authors use word processing software. They didn't use typewriters. The editorial and production workflows were digital. Today we consume lots of our content on the web and we do a lot of reading on phones and tablets and apps. E-books are commonplace. Why this is so fundamental is that Assistive Technology benefits from digital files, digital system, digital devices and the accessibility of those things benefits from interoperability and open standards which are really key to where we are today.

Web technology is really fundamental. The whole publishing technology tool chain is based on web technology. The markup is XML and HTML. CSS is used for rendering. There's standards you may not be familiar with called aria which is used for structural semantics. There's MathML for math. Tables are increasingly coded as HTML tables. Why this is so important is for a couple of reasons. All of those technologies are free, open interoperable W3C standards. All publishers and their vendors and suppliers all use those technologies. They all know how to use them every day. They are now the basis for accessible publications. EPUB is built on web standards. George will go through the history of how we got to that point because way back when that was the vision.

HTML markup is the basis for the mark up and Assistive Technology now is built to understand HTML mark up.

I mentioned that CSS is significant. The technology that specifies the presentation of the content separately from the mark up in the content itself. In other words, it is based on the content markup but the content markup instead of saying this thing should be bold it says this thing is a level 1 heading, et cetera.

And then that HTML markup has been supplemented by aria markup which are one of the things that aria does is provide roles for content objects. For example, HTML uses sections. Aria enables you to specify the section is a chapter. Aria lets you say there's a foot note et cetera. MathML and HTML tables are understood by Assistive Technology.

Again, as I said before, virtually all publishers and vendors and suppliers that they work with all use these technologies.

So what does this mean? In today's workflows publications can be born accessible. That was claimed by Betsy [inaudible] at benetech. She was hearing the phrase born digital and she thought if it's born digital, it should be born accessible. Accessible EPUB 3s can be a standard product of workflows.

Here's the catch: You have to use this technology properly. EPUB 3 doesn't guarantee accessibility. It enables accessibility. For several years now I have been trying to get traction to get accessibilityability. The ability to be made accessible. Some of the things that are fundamental to get them right is you have to use the HTML mark up properly. The H1s have to be nested properly. The content has to be in logical reading order. It must have structure to enable navigation. You need to be able to jump to chapter 5 if you want to start reading at chapter 5. You need to be able to skip over content. If there's a side bar you don't want to see, or you don't want to read the foot notes you need to be able to skip those. And print disabled user depends on that. The big one that's fundamental is images need description. Those are not rocket science. Most of the vendors can do most of those things. The trickiest of them is the image description. Increasingly it's become easier to get those right.

Ideally, accessibility should just work. The publications, the systems, the devices that we are all using use web technology. Accessibility is fundamental. So you should be able to expect those things to be accessible.

So accessible EPUB 3s can be a standard product of workflows. Websites and publications can be navigated by keyboard access. Someone who can't touch the screen can navigate through the document properly.

Phones and other devices are just built now to read text aloud. Every iPhone in existence can read text to you. Even videos often have closed captions with them. Those things are important. Much of that technology is built by the accessibility community. They make things better for everyone. I will turn this over to Luc. You can tell us how we got here and where we are now.

>> Luc: How did we come to this EPUB ecosystem? It all started with international digital [inaudible] in 2006 which published the specification. Previously there was an open EPUB [inaudible] that was considered. EPUB was there and used for several years and it was a very first interesting step to produce e-books using web technology. These technologies were available during that time HTML 4 and CSS 2 and even those technology were not fully deployed in EPUB 2. There were restrictions. The big effort was done in 2010 and 2011 with DAISY consortium. Markus was the CPO of IDPF. There was a collaboration of DAISY consortium and IDPF all over the world. It was an advancing technology in EPUB 3.0 was published in 2011. It was the best EPUB because it was using the latest technologies in the web, HTML 5, CSS 3 and also many unicorn. Meaning a new language of the world could be used inside EPUB. It was very interesting for those countries with specific scripts like in Japan. EPUB 3 was marketed primarily for the possibilities of multimedia and interactive activity. It was not the main speech. Obviously it was much better than EPUB 2 in even what EPUB 2 was able to do text and simple books.

Then it really was interesting for trade because it was inside the specification of EPUB 3 was accessibility. It was accessibility enabling in EPUB 3.

So the years went by and by. There was a lot of effort to make EPUB 3 adapted. It was not adapted at the beginning. There were initiatives like the [inaudible] foundation. There was tests of EPUB 3.5 available. There was anticipation for the future inside IDPF and we worked on EPUB 3.1. This EPUB 3.1 was tentatively improvement on web technologies to go further. It was never adapted. There was no one adopting 3.1. Particularly there was an issue with very interesting free tool EPUB check who help us to say that if an EPUB file is really conformant to the specification. EPUB check was left in the air in 2016. There was no development and it was not addressing the EPUB 3.1. We were in an uncomfortable position at that moment.

We wanted to move EPUB 3 -- the publishing industry. There was a movement that is explained in the next slide. There was a movement in IDPF to come near the web technologies and movement of merging inside W3C was initiated. That was evolved in 2017. It was then some label created that was publishing at W3C with several groups. It was -- the movement was interesting for EPUB and the publishing community because it helped us to be at the very heart of where technologies are written and forged.

This was very interesting and there was several groups created. There was publishing working group for the future starters and that's not our webinar for today. For the maintenance of EPUB 3 which was already existing, a community group was created, and this community group took in charge this EPUB 3 specification. It was free membership group. It is very active. There are more than 250 people there.

This active group did a tremendous job to maintain EPUB 3 specification. There was another group who was there to look at the global industry and make the EPUB 3 ecosystem in a better position.

What was the goal of the publishing business group? It was to put back EPUB 3 at its highest value. Meaning not just leaving the uncomfortable situation and make it more stable and future proof. Another goal was then obviously to tell the publishing industry that EPUB 2 was obsolete somehow because even in the simple and full capacity of EPUB 2 -- not speaking of [inaudible] -- EPUB 3 did better. Particularly because EPUB 3 was accessible by nature and in specification.

So what did the publishing industry for EPUB 3 in these last years until now? There was a revision of the specification which was done by the EPUB 3 community group. This revision was done in this large group with a lot of consensus and revisited the version of EPUB specification so that it comes to a new release which is fully compatible with 3.0. There's no compatibility issue. In the same time, this new release want to be future proof. Meaning they built it on the latest HTML and CSS. Not this specific version but the latest version that of HTML and CSS that W3C will ever publish. This was successful in May 2019. It was published as a final report and the name of it is EPUB 3.2. This is the best ever EPUB specification we ever had.

The publishing business group worked for the ecosystem globally. First we started a revision of EPUB check because as I said since 2016 nothing was done. So we launched a 2 years plan for development of EPUB check to fix it and upgrade it. This plan is not finished. It will be finished this year in 2020. There is still some actions of development like test suite, API, documentation. The DAISY consortium is working on it still now. The publishing business group also in parallel initiated this international fund raising so we can make this revision of EPUBcheck fully developed.

This group has raised almost the full money we need. I use this webinar to call once more for help because as you can see we still miss $10,000K dollars. So the link encourage you to click and donate.

These two actions are important for the ecosystem. We have specifications that are future proof. We have EPUB 3. [inaudible] which is a diversity document. This EPUB accessibility is the specification to explain exactly what is the goal of accessibility in EPUB 3 and you have these accessibility techniques document that plains precisely what needs to be done with tags, HTML attributes. It's easy to main how to build EPUB accessibility document.

So now we have these two free tools that are up to date. We have EPUB checker accessibility. Any [inaudible]. EPUB check is fundamental. This ace tool help us to be sure that the techniques that bring accessibility inside EPUB files are correctly used. This tools enabled publishers to be clear with that and also their suppliers. So the full production workflow is okay.

So trade publishers have been trying to start with accessibility. Many publishers over the world are doing their best and doing real production like bill explained with his workflows. I can explain where I spent Hachette Livre decided that in 2018 any new titles published in EPUB 3 should be born accessible. Meaning that the workflow has put ace as one of the checking tools besides EPUB check and the others. Before any new titles in EPUB 2. For that effort we won an international excellence award. In 2019 we signed the accessibility book [inaudible]. Just before I left in March 2020 after producing every new title we also managed to push for each of these the metadata. There are many examples all over the world. This is one example. This is to show it's possible to go this route. I give the [inaudible] to George who will explain about accessibility inside EPUB.

>> George: Thank you. I will talk about baking in accessibility because it's not frosting that's spread on after the fact. We mean we are talking about integrating accessibility into all aspects of publishing. I'm going to talk about how we did that.

We are going to look at the various domains. I'm not going to go through the rest of this slide because the other slides address each of these domains.

Next slide.

So first of all, within the W3C the web accessibility initiative which brings you great things like WCAG2.0 and now 2.1 which I think most people are aware of certainly for the web and it also applies to publishing. Publishing is based on HTML, CSS and these things and we can use WCAG with EPUB files.

So the DAISY staff are assigned to work on a wide range of activities within the W3C and within publishing in the W3C.

So Avneesh is our COO for DAISY was elected to the advisory board for the W3C.

Matt, myself and Avneesh work on the WCAG. Romaine is working on validation at W3C. And he see also the lead coder for EPUB check and one of the lead coders for Ace. Other DAISY consortium members like benetech are -- have people assigned to various working groups and Charles is one of those great contributors.

I've been chairing the steering council of the web accessibility initiative since it began.

Next slide.

So the legal framework, the convention on the rights of persons with disabilities we worked with them and made sure that the issues surrounding publishing and accessibility were baked into that international treaty. We know that ADA and 508 apply to the web and we have been working on pushing that into publishing and now 508 requires accessible office documents. Now the EU directive which is coming into effect in the years between 2022-2025 have a clear mandate to make sure published materials are accessible.

Next slide.

So we focused on standards from the very beginning.

People from the blindness community integrated in with all of the digital publishing that started in 1999. We engaged technically and integrated with the oversight and officers in these various organizations. So the whole business was integrated, not only technical things.

We brought in technical expert like Markus who was mentioned who was the CTO for the DAISY consortium. And he also became the CTO of IDPF, international digital publishing forum.

So you get this baking in aspect where the best people in accessibility are leading developments in publishing.

We initiated the EPUB accessibility conformance and discovery specification within the IDPF and now in the W3C. Thank you to Google for funding. Google impact challenge helped us pay for these things.

And we have moving this back into ISO and right now it's voted on in the United States by NISO and if you haven't voted yet and you are a member of NISO, get out and vote.

And Matt from DAISY is the editor of the various specifications. He see an expert. Whenever anything is put in a specification, it's put in by him. He knows accessibility inside and out.

Authoring and publishing ace by DAISY which was mentioned. That does the automatic checking of the book, but you can only do so much with automatic checking. So we created the simple manual accessibility reporting tool. Which is like a check list on steroids. Ace generates reports that gets pulled into smart and you can manually check the things you need to do. There's data visualizations to make it easier to check. All of the images are shown along with their alt text and surrounding material; there's a listing of the accessibility metadata. The DAISY knowledge base is linked from all of these tools to the knowledge base which is maintained by Matt and the rest of us contribute and review to make sure that the best practice for accessibility are front and center and easy to find.

Conformance and discovery metadata. Inside is metadata. It points to WCAG2.0. We didn't reinvent anything. We just point to what's there. We also added requirements that are specific to publishing that the web really don't know anything about page numbers for example, foot notes. They don't exist on the web, but they exist in publishing.

Conformance metadata points to the WCAG and EPUB. Certification metadata. And discovery metadata that is intended to be exposed in catalogs and bookstores when people buy so that people can buy born accessible. So if the metadata is there, you can see it's born accessible and that's the one you want to buy.

We have reading systems and apps that are being tested. We test for accessibility.

A whole lot of different Assistive Technologies are used to test the various reading apps and systems on all platforms and different disability groups. We provide feedback so they can improve their reading system. The next time there's a release we are notified, and we test. If they have questions, they come to us for help.

We make public information available about all the apps that are on inclusive publishing.org. We do a roundup of all the different reading systems and how well they work with different technologies.

When we do our testing, we identify issues and sometimes it's not with the reading system but with the Assistive Technology itself.

So we can provide feedback to them and let them know how to improve their products. This improves the Assistive Technology as well.

Next slide.

So distribution outlets and libraries. We are producing -- this is in the community group at W3C which is free to join -- using experience guide. This metadata is not intended for human consumption but it's machine readable. So it gets translated. We want to see all the bookstores and libraries expose this accessibility metadata. It supports the buy born accessible. We want everybody to be buying born accessible. And lastly, in higher education we got an EPUB working group. We are looking at getting higher education adopting EPUB. The publishers are there. Literally all of it top publishers are producing born accessible EPUB 3. They are working on getting certified by benetech’ s global certified accessible program. Many of the publishers are doing that now.

We are doing a whole bunch of webinars in the higher education space about EPUB and we are in the process of planning next year's sessions right now.

We did identify barriers and then we set out to address those.

We now have word to EPUB, and all the other word processors produce EPUB and we'd like to see professors producing EPUB themselves and put those along with the PDFs. We need to get this initiative going in K-12.

So what we'd like to see is a whole society that's inclusive and embraces accessibility.

We know that access to information is a civil right. In the access age this is a human right. The publish industry has been terrific. It's being woven into their workflows and baking it in right from the very beginning.

So, I would love to see teachers in middle school when we start teaching writing for the students to be creating born accessible documents from the very beginning. It's not that hard. So I'm going to wrap there.

>> Richard: Thank you so much bill, Luc and George. Bill, you have the control if you would move us to the contact slide. That would be wonderful.

We are going to cover as many questions as we can. Please note down the contact information. This will be posted on the webinar page in the next few days. [Reading from PowerPoint].

So they are there to take your questions. We have time for questions now. Let's get to that. One of the questions we have was really about -- Bill, you talked about the transition from the old fashion ways of doing things to digital. When would you say that transition really took place? Within the last year or so or longer where content was being produced digitally and consumed digitally?

>> Bill: It's been happening for years. It's been a very gradual evolution. It's accelerated in the past couple years. I usually attribute this to the fact that the standards all fell into alignment. Global accessibility standards are basically all based on WCAG and the other W3C accessibility standards. They used to differ. They differ much less than they used to. The tools and technologies that not just publishers use but suppliers and vendors use are the same standards that are required for accessibility. One quick comment in terms of something that accelerated this development is the influence of important publishers. Kudos to Luc because by doing what he did with Hachette Livre they made all of their vendors had to deliver born accessible EPUB 3s. Many people think the publishers make the EPUB 3. Many of the vendors do that. They trained the vendors. So there's lots of vendors out there in the ecosystem that know how to do this properly. The same thing with higher education. Another one I will highlight is macmillan learning. They were the winner this year of the inclusive publishing award. They spent a lot of time working with their vendors and now those vendors are doing that same work for other people.

>> George: I'm going to jump in and say with the word to EPUB tool anybody -- independent authors, professors -- can produce born accessible documents. They need to follow normal heading structure, name styles in Word and use the functions that are there to create foot notes and citations and that automatically comes out as accessible. And quite frankly now that that tool is out, I think that's going to hopefully put pressure on pages and Google docks to up the ante a little bit.

>> Richard: While you have the mic.

>> Bill: DAISY did that.

>> Richard: A couple questions for George. While we heard a lot about the industry and technology, could you set out for us what difference this can actually make to a reader or a learner who has a print disability? What impact does this technology mean for them? The link question to that is how does a librarian or someone working in university tell whether an EPUB is an accessible EPUB?

>> George: Okay. First of all, the world line union has had this chant: We want the same book at the same time at the same price. Now with the publishers producing accessible materials and in higher education the publishers I'm seeing are putting in extended descriptions. So complex graph are fully described. There's no need to go to a service -- your disabled student service office or to an organization that has served BD which is separate and apart. They can just buy the book and start reading it.

The second question is that the same answer applies to libraries and bookstores. For example, VitalSource when they ingest the EPUB, the accessibility metadata is inside, baked in to the EPUB. The accessibility summary that is a human readable description of the accessibility features of this book conforms to certified are right there. They can extract that when they ingest it and make it available on their website. They have been doing this for quite a while. The publishers are producing onyx metadata about the accessibility of the book as well. Code page 196. So librarians need to put -- and professors -- need to put policies in place that we will buy born accessible, we demand born accessible. If we don't see the accessibility metadata, we are not going to buy it. They need to expose that in their library catalogs.

>> Richard: A question for Luc. As someone who was working in the heart of the publishing industry, could you speak to the question of the third-party certification and maybe also self-certification? What’s the importance to give buyers, end users, customers titles about accessibility?

>> Luc: The first type of certification is almost at the heart of the publishers production. Publishers have to make their suppliers produce born accessible using the techniques and the free tools to check like accessibility checker for EPUB. Obviously the publishers also have to work with their authors to make image descriptions available. That's something that really at the heart of editorial process with the question of context of this image in this book.

Then all this is done by the publisher. The first step of making the information available that this EPUB has been built using the best knowledge and tools to make it born accessible is the first step of information.

This is not a certification. It's a note to declaration. It's very important to have this. We don't have it right now for many publishers, many suppliers and websites. The websites and for librarians and to buy e-books don't show yet this information. That's very important. This is also inside the European accessibility act that will enforce all the step of the supply chain this information to be visible about accessibility and the website will have to display that when the European accessibility act will be enforced in 2025.

So every piece of this has to be ready. Obviously there are some high-level certification process by some suppliers like benetech. In that case it's really high-level certification and you may have a very good stamp and you are sure. This is a continuous process of information. For many books and I know the trade part and less the [inaudible] part.

For the trade part auto declaration by the EPUB publishers is very interesting first step to enable visually impaired people to find books as soon as their published in digital format with a large quality of accessibility. This is the end of the book famine. We are on it now.

>> George: Can I add a little bit to that? Whenever a publisher is producing something, they making the assertion that this is accessible. Third party helps them with areas of problems. With the ace that's out now we have the ability to edit the accessibility metadata inside the EPUB so that's a nice new feature.

>> Richard: We are going to squeeze in one or two questions. This one is to Bill because you work with a range of vendors and publishers. This comes from Alice. She works with students. When she was looking for titles on the various platforms she wasn't finding the titles that the students needed in EPUB. What's your crystal ball for how many books are available in EPUB and when will all books be available in EPUB that the students need?

>> Bill: First of all, I don't know whether she see referring to higher education or K-12. Higher education will come sooner because of the formatting and the visual aspects that are so fundamental. Often times there are EPUBs available but the campuses, the disability services offices, et cetera, even the students asking for accessible content are not yet completely familiar with how to use EPUB or the fact that EPUB is so inherently accessible and so much better than the formats that they are used to. I'm in a melon funded project with academic libraries, disability services offices and bookshare, Internet archive and [inaudible] to address this problem because a lot of it at this point is education more than it is the technology. The technology is pretty much there but the education is lagging.

>> Richard: Thank you for that. The last question which we will just squeeze in before we wind up is from page. Page is talking about working as a librarian and the importance of open educational resources. We have talked about publishers and vendors what's the relevance of this journey to open educational resources?

>> George: Most of the OER out there is not using EPUB today. However, they -- the tools that they are using to produce the OER like Microsoft Word or Google docks do support EPUB. I think we need to see a push in OER to use EPUB as the format. Also to include the accessibility metadata because the publishers in higher education are really having their toes held to the fire to produce born accessible grade materials. The OER needs to meet those same high standards.

>> Richard: Thank you for that. Okay. We are coming to the end of the session now. Thank you to everyone who joined us for today’s webinar. Bill, Luc and George, thank you for sharing your wonderful insights and information.

Just to remind you, this was the first of two webinars examining the revolution of born accessible digital publications. On June 3rd part two will consider "The future of accessible publishing and standards “where are we going? Coming up in the next few weeks we are delighted to bring you representatives from Amazon, Google, Adobe and Apple, together with other esteemed panelists from a fantastic variety of organizations:

May 13: Easy access to books and articles through a smart speaker

May 20: Leveraging InDesign for Accessible EPUB Creation

May 27: Accessibility at Apple

Find out more information at daisy.org/webinars, where you can also sign up to the webinar announcement mailing list to learn about new topics as we add them. If you would like to suggest a subject, or if you are considering presenting a webinar, then please email us at webinars@daisy.org

Thank you for all of your questions. I hope you will join us again next week. In the meantime, thank you for your time and have a wonderful rest of your day. Good bye.