PDF Remediation Software Discussion - Event Transcript

Opening

Lori Litz

Good day, everyone. Welcome to today's Accessibility.com event, PDF Remediation Software. So we're picking up from where we left off last month where we talked about PDF remediation services and now we're going to dive into some of the companies that provide software for those of you that maybe have tons of documents to do and you want to do them yourself, or it's just something you want to add to your toolbox as you travel along your accessibility journey.

My name is Lori. I'm the Director of Conferences here at Accessibility.com, and we're so excited to have you here again with us today. Our event starts off with a PDF Remediation Software discussion with our friend, Casey. Our President, Mark, and Casey will talk about some of the things you should look for and what PDF remediation software should do.

And then through that event we will have our sponsors, Crawford Technologies, PDFix and DocAxess. We'll have some deep dives into those into those companies and what services they provide and what they feel sets themselves apart from the competition. Then we'll go into audience questions and answers. Please type any questions you have for that panel into the Q&A.

Let's have the chat be for discussion amongst yourselves and so on. But the Q&A section is where our team is looking. If we don't get to all of the questions today, someone from today's event will follow up with you and answer that question for you. After this first part today, the PDF Remediation Software discussion, there will be a presentation at 130 Eastern time from Crawford Technologies called Document Remediation The Easy Way.

At 2:00, PDFix will do Trillions PDFs to Make Accessible. And at 2:30, DocAxess will have PDF Remediation at Scale: Debunking Myths and Unveiling Realities. So you have a full day today. All these events are recorded and will be available this evening for you to access on demand. Once you register for this event, you'll receive an email from me with instructions on how to access that later tonight.

Without further ado, let's head on over to PDF Remediation Software with Mark and Casey.

Presentation

Mark Shapiro
I'd like to thank Casey Naiduk for joining us. Casey is a leading expert in digital accessibility. He's worked with Wells Fargo, Oracle and he's currently the Senior Manager of Digital Accessibility at Bristol Myers Squibb. Casey, thanks for joining us today.

Casey Naiduk
Oh, it's my pleasure. Happy to be here, Mark.

Mark Shapiro
Thanks. For starters, what should companies expect from PDF remediation software?

Casey Naiduk
That's a, that's a simple question with a probably nuanced answer. So there are a number of different PDF remediation softwares out there. Some of them are more of like a guided, guided approach to more of a manual effort on your part and some of them are leveraging more of that automation or or AI. So they can do quite a bit.

They're capable of quite a bit. But generally what a PDF remediation software will do is you upload or access an existing PDF document and it will help, at the end result, help it be tagged for accessibility. So, so the general, the general goal for PDF remediation software is for the final output to be an accessible PDF document that would be consumable and accessible by by everyone.

The primary factor there is to make sure that the document is tagged for accessibility, which is just the process by which semantic structure is added to a PDF document to make it primarily readable by screen reader users.

Mark Shapiro
So why would you use PDF remediation software, not just an accessibility, a PDF accessibility checker?

Casey Naiduk
Well, a PDF accessibility checker can be a great tool to tell you how how well you've hit the mark, at least to the extent that those tools are able to, you know, identify how well you passed their rule, that they're checking against. The benefits of using a software... Well, there are a few. One is to help you with the actual tagging, the actual remediation of the document.

So a checker is kind of like, you know, a status check. How'd you do? How accessible is your document? You're passing these checkpoints, these checkpoints. You're failing for those checkpoints. But the remediation tool can help you actually fix the document so it's usable and accessible.

Mark Shapiro
How does remediation software handle items that should be manually reviewed, such as things like reading order or tables or color contrast?

Casey Naiduk
Sure. So again, they vary a little bit in the approach that they'll take, but typically think, well, let me start by saying this. Anything generated by automation such as, you know, a PDF remediation tool that's using any any form of AI or generating some output for you. The burden and the onus is really ultimately always on you to kind of verify that it did its job.

Right? So computers are really good at a lot of things, but some elements require kind of interpretation or validation. So it will always be recommended regardless of what tool you're using to verify that it did the job correctly. Now for some of your examples, like a table, that's a, that's actually an area where PDF remediation software can be really helpful, especially for complex tables, where you need to add things like column headers, row headers.

Sometimes there's nesting involved there. Those can be tedious and tricky and very manual to do and time consuming to do from scratch. And a lot of times, at least in my experience, what the PDF remediation softwares will do is they may not always get it perfect, but they'll at least give you a really good starting point if you do need to fine tune it.

And so the key benefits there are speed, oftentimes accuracy, and then depending on your own accessibility expertise as the user of that product, you know, let's let's say you're you're really not experienced in this area. Well, it can do a lot of the heavy lifting for you and more likely than not, get it right. And you may be okay there. If you're somebody like myself or yourself, again, it'll give you that starting point where you may want to fine tune. But it's it's a lot easier than starting with a blank canvas.

Mark Shapiro
In terms of the companies are out there, there's a there are a lot of companies that offer PDF remediation software. What should people generally look for when they're, you know, deciding on one vendor over another?

Casey Naiduk
Well, that's a great question. There's a lot of factors at play. Budget is one. Right? How much do you want to spend? Speed is another. How quickly do you need documents turned around? But in general, the way that I, I would judge a PDF remediation software vendor is similar to how I would judge any accessibility vendor or any vendor in general, which is, you know, you want to make sure that you have kind of a support model in place.

And so when you're using any product, whether it's software as a service or anything, there will be hiccups. There will be questions. There will be times where you kind of need that support. And so for me, I always, I'd always opt for, if I had the choice, a company that has kind of a robust and a transparent support model.

Now, when you're choosing, you know, PDF remediation. Uh, software vendor in particular, some things you might want to keep in mind are volume. Right? So make sure that the that the license that you're getting or the product that you're purchasing that it can handle the type of volume you're looking for. Right? So there may be a it may just be a subscription model with, you know, unlimited documents.

It may be a per document or per page type of thing. Make sure that it fits. It fits what you're looking for.

Mark Shapiro
Thank you, Casey, for taking the time to share your insights about the PDF remediation industry. Next, let's do a deep dive into three leading PDF remediation software companies. We're pleased to have with us today Doug Koppenheffer, the VP of Accessibility Center of Excellence at Crawford Technologies. He's going to let us do a deeper dig into their remediation software.

Doug, thanks for joining us.

Doug Koppenhofer
Thanks for having me, Mark. Very excited to be here.

Mark Shapiro
To start with, can you tell us about your company and how it how it offers its PDF remediation service software?

Doug Koppenhofer
Yeah, absolutely. So I'd like to start by saying that our CEO and Founder is still the CEO of the company, Ernie Crawford. We got started because Ernie realized that he his company could create accessible PDF by virtue of doing transformation of documents into PDF format.

And at the same time, his mom developed some visual disabilities and he decided, you know, he had to go up and read his mom's financial statements and tax documents to her and realized, hey, there's a there's a place for accessibility and we should pursue this. So what what we do and what Ernie has developed is the ability to make document content accessible on premise with software and also off premise through a remediation services approach that we have in a secure facility both in Toronto and in Potsdam, New York.

So that's the high level on how that works.

Mark Shapiro
What's a typical use case for your customers?

Doug Koppenhofer

They range from very automated kinds of use cases where you have a set type of documents. Let's say it's in the health space. Let's say it's Medicare Advantage programs. They generate the same kinds of documents for their members over and over again. And there are 41 types, very specific types of content that go out.

We pre, we establish the tagging rules for those documents in advance and actually get the CMS models for those new updates each year. And we update that automation every year so that you're ready to roll. So that's one use case. Another one is much more public where you have public publicly discoverable content on a website. You know, it might have been generated by any number of authors.

And in that case, what we do is we scan the site and find the files that are on the site. We test them for accessibility and then, you know, ultimately put them into that workflow. So there's a couple of use cases that we see quite often.

Mark Shapiro
How do you charge for these for the software?

Doug Koppenhofer
A lot of different ways.

I mean, we run the software either on premise so you can license the software based on the number of pages that you process on an annual basis. There's also a way of tying it to the total number of clients you have. There are a lot of different ways of licensing it on premise. And then, you know, in a cloud setting, we have a secure cloud product.

And that really is based on the number of pages that you pump through it.

Mark Shapiro
What is it that really sets you apart from other PDF remediation software companies?

Doug Koppenhofer
Yeah, I think it's, I think it's sort of the combination of capabilities, right? It's the ability to do the high volume stuff that can be templated. It's the, it's the using the AI for the ad hoc or static content and the ability to do remediation as a service and just the expertise of the company, you know, in terms of being able to, you know, know exactly what what WCAG is doing next.

Mark Shapiro
Very cool. Well, Doug, thank you. Appreciate you letting us do a deep dive in your company.

Doug Koppenhofer
Thank you very much, Mark.

Mark Shapiro

We're pleased to have with us Jozef Baranec, the CEO from PDFix who is going to allow us to do a deeper dive into his PDF remediation company. Jozef, thanks for joining us today.

Jozef Baranec
Yeah. Hi, everyone. We are a company which is based in Slovakia and Europe. We have been working with the PDF for the last 25 years. The, our main goal was really to make work with PDF much easier for everyone. So our small team with eight people is dedicated to really build tools and revolutionize the document manipulations and use the latest AI technologies and the recognition algorithms to unlock the full potential of PDF documents, making them accessible and user-friendly for everyone.

Mark Shapiro
What should a company budget for your software?

Jozef Baranec
I could say that we can, we can provide a solution for almost any company size, which means that the budget should not be a limit for anyone. And basically we are trying to find the solution for any kind of a customer, even for the low cost and also for the high volume expansive solutions.

Mark Shapiro
What sets you apart from other PDF remediation companies?

Jozef Baranec
With so many years of PDF development, we realize that automation will be something that will be crucial with the given the amount of the PDF documents in the world. It's huge. And no manual work is possible to cover all solutions. So that's why from the very beginning we focused on automation and also portability and cross-platform development.

Mark Shapiro
So if a company has, say, 100,000 documents, they could actually do all this on their own with your software? They could have it be remediated?

Jozef Baranec
That is right. As I said earlier, it requires a lot of investigation of the document types trying to find the metrics in the files that allows us to really prototype or process the documents in a way that the quality that will be achieved after processing will be best possible. And of course, we have we have to generate fully compliant PDF UA documents or WCAG documents.

Mark Shapiro
Okay, Jozef, let's say you're working with the university and they have documents that are dating back from the 90's as PDF documents. How would your software handle remediating those?

Jozef Baranec
Well, the process can be split in two, of course to the analyzing part, which will identify the documents, classify them based on the type, based on the structure layout and so on. And then figure out, figure out what would be the best approach in autotagging the document, the processing the documents. Once we have these in place, we can set up a set of tools that will automate all the processes and specify only that portion of the work that must be done manually to minimize the amount of effort that is necessary to process.

Mark Shapiro
Our third company deep dive today is with Moise Akbaraly, the Co-Founder of DocAxess.

Can you share with us the background of your company?

Moise Akbaraly
Yes. Well, DocAxess has been founded in 2020. It's originally a spinoff of a previous company called IPEDIS, which is specialized in digital consulting services. So we've been, we've been doing accessibility for 18 years now. It's a tech company composed of engineers, PhDs and developing technology for PDF remediation purpose only. We only doing PDF remediation.

Mark Shapiro
So when you say you're doing PDF remediation software, are people... How are people getting you the documents? What are you actually... are they sending you the documents, you're remediating and sending back? Or is there? What's the process?

Moise Akbaraly
Well, the process is very straightforward. We've built a range of software solutions, a SAAS platform is available for, for our client where they can drag and drop their, their documents and, and our AI system will process all the, all the work and then all the, all the documents will be available on the platform.

The other way is to connect to our API system and is meant for large clients and large corporates that's processing quite sustainable amount of PDFs throughout day. Well, the one of the key thing of our system is our AI doesn't really care about the complexity of the PDFs you're sending to us. We've built a technology that can deal with any type of PDF, whatever the complexity you may have. The

the only thing is we, we processing with what technology native PDF.

Mark Shapiro
What really sets you apart from other remediation software companies?

Moise Akbaraly

Well, we have a very tech approach to, to the problem. So it means that we investing a lot in modeling and AI systems. I said previously that well, we've automated today 80% of the work and in some cases we were able to automate 100% of the work.

If you have, let's say, an invoice or a payslip, for instance. So what sets us apart is what we delivering by the end of the day is 100% compliant document according to PDF/UA Standards.

Mark Shapiro
Thank you for joining us today. Now let's answer some questions from the audience. We have quite a few here in queue, so if we don't get to your question today, our team will follow up with you. The first question How do you see AI transforming PDF remediation software? Doug, why don't you take this one?

Doug Koppenhofer
Yeah, So that's a really... it's interesting.

This year at CSUN it was all about AI. The previous year, we had maybe two or three questions that came up. But it's all about AI these days. And I think our approach to this is that it's very much a trust but verify model. We definitely use AI. I mean, we use it when we're setting up the automation templates for high volume content, but then we about validate that what we've set up actually is going to generate good results.

Same thing in the static world, we use AI to automatically tag the files, but there's always a step at the end where you have a person with a workstation that just runs through and makes sure that you know that the the alternate text is correct, that the the read order is good. So it's it's helping time to market immensely, we think.

And we actually have you know, we have real results right now where we're seeing 5 to 8 times faster time to market through the AI components. But it's always, you know, we have to validate at the end. That's what we're seeing.

Mark Shapiro
Jozef?

Jozef Baranec
It's an interesting topic. And we also use AI in our solutions more and more. Of course, it has some limits at the moment, but what I can mention is that it really can help in the most outstanding issues. Issues, for instance, generation of image captions or identification of the language. These are the topics where it usually also our software can help and really simplifies the work.

Mark Shapiro
Moise? Anything to add?

Moise Akbaraly
If, I mean, if we look for the past ten years and even more, I mean we've been doing PDF remediation manually, which is very, very time consuming. So AI is is is a milestone in the, on a journey because it's now we can, we can consider moving PDF at scale, meaning that in the previous days you had to select which one of the PDFs you wanted to remediate because of the cost, because of the time it takes to deliver.

And now you can, you can clearly imagine that tomorrow you can deliver every PDF that you put in line at a very cost effective way and at the speed of the light.

Mark Shapiro
All right, next question. Is there anything that remediation software cannot remediate? Doug, why don't you take this one?

Doug Koppenhofer
Yeah, I think there are some components that are very difficult.

So one example is in the in the high volume space, you can generate a PDF file that has missing Unicode values on the fonts. I know I'm diving into the technical weeds here, but you can actually generate a file that presents beautifully in PDF and is sort of garbage underneath and doesn't have the textual information that you need for a screen reader to interpret and navigate it.

So there are some tricks that we can use to interpret or interpolate with the fonts should be, or what the value should be, But that's sort of a danger area. If you generate, you know, files without those Unicode values on the fonts or the characters, it really is tricky. So that's one to watch out for. I'm tempted to also say image content, but there are ways around that.

I mean, you can you can OCR that content pretty reliably. It's just that you add a step to that because you have to validate the OCR that you generate and make sure that is actually correct before you even tag it and make it an accessible version. One interesting thing that we're seeing a lot of lately too, is the ability to OCR, keep the image exactly as it was, and then put the the, the text behind the image so that you actually see the original image.

What you see, we have a screen reader version that's sort of underneath the curtain, as it were, that you can use to, to read aloud the information. So some tricks in the industry there.

Mark Shapiro
Cool. On average, how much time does PDF remediation software save me over remediating it myself in Adobe? Moise, why don't you take this one.

Moise Akbaraly

Yeah well if you if you could take consider a hundred pages document if you do it manually it will take you about 4 to 5 days to remediate Whereas if we... if you upload the the PDF on the system it will take, it would be, it would be delivered within the day thanks to AI.

Mark Shapiro
Is there a standard quote tagging or tagging tree format that I should follow? Where is this published if it is?

Casey Naiduk
Well, there are a lot of great, there are a lot of great references on the web for for learning about tagging and tagging structures in general. Adobe, right? In support of their Acrobat Pro product, they have a lot of great and freely accessible resources on the web that can kind of explain basic tagging and tagging structure to you.

There are some some government sites. I think HHS, Health and Human Services, that's what HSS stands for. They have a great document out there as well that will help you learn the tags required for for conformance. In that case for section 508 conformance. But they're easy to find.

Mark Shapiro
How does PDF remediation software confirm that a screen reader can accurately read the document? Jozef, why don't you to take this one.

Jozef Baranec
Our software is using veriPDF as an integrated validation tool which can be used directly on an open document, and you instantly see the problems regarding the reading order and correct reading in the screen readers. We, our software is using a called so-called derivation algorithm which transforms a PDF into readable HTML website where the user can check the order and see how it performs in the deconstructed feed.

Mark Shapiro
Okay, great. How does remediation software handle tables? Moise, why don't you take this one?

Moise Akbaraly
Yeah, we we managed to crack this this main part with modelize and automating very complex tables because we've been we've been doing we've been delivering financial reports for the past 18 years. We've managed to, thanks to to those archives and to thanks to the labs we're working with here in Paris to to to modelize those complex tables and help delivering a very consistent content for the screen readers,

Mark Shapiro
Does PDF Remediation software, correct formatting errors, spelling errors and grammatical errors? Casey? Why don't you take this one.

Casey Naiduk
To my knowledge, no. So a PDF remediation software primarily is focused on the tag or tagging tree, and that truly is a behind the scenes thing. That's, that's not visible unless you happen to have the tagging tree up. So typos, grammar, things like that... To my knowledge, I don't know of a PDF remediation product that would cover those.

Mark Shapiro
All right. And that's all the time we have for today's questions. If we didn't get to your question, we'll follow up. Thank you. Casey, Doug, Jozef and Moise for sharing your time and expertise with us. And thank you for everyone in the audience for being here today.

Closing

Lori Litz
Thank you, Moise, Jozef, Doug, Casey and Mark for that great presentation on PDF Remediation Software. As I mentioned at the top of this discussion, this event is recorded and will be available for access on demand. Later this evening you'll receive an email from me with instructions on how to access that. Stick around. There's three more presentations today that you can access from the lobby.

Next up is Crawford Technologies. With document remediation the easy way. Then at 2:00 trillions PDFs to make accessible by PDF Fix and at 230 PDF Remediation at scale. Debunking Myths and Unveiling Realities. All of the presentations will be available on demand this evening. So if you have to step away or you have to take off, no worries, you can.

Once you get that email, you can access the on demand content and watch at your convenience. Thank you all so much for being here today. We will be back in May with three events. The first one is Accessibility Services for Small to Medium Businesses. That's on May 1st. May 14th is Accessibility Services for Enterprises. And on May 30th, we have Accessibility Services for Mobile Applications.

So some very exciting topics coming next month and more to come after that. In addition to watching the presentations from today's sponsors that are here, Crawford Technologies, PDFix and DocAxess, you can also head out to the Virtual Expo. You can access that from the lobby and you can step into the booth and visit with the companies.
You never know what you might find when you go out there and visit with them and get some more information on PDF remediation software. Thank you again for being here today. Have a great rest of your day.