

PDF Remediation at Scale: Debunking Myths and Unveiling Realities Breakout Session Transcript

Moise Akbaraly

Hi everyone. And welcome to this presentation of DocAxess regarding PDF Remediation at Scale: Debunking Myths and Unveiling Realities. I am Moise Akbaraly and I'm the Co-Founder of DocAxess. I've been doing... I've been doing PDF Accessibility and PDF remediation for the past 18 years. I set up with my partner, Jackir, DocAxess in 2020.

And before, before that, we've been doing PDF remediation with a previous company called IPEDIA which is a, accessibility consulting services agency. So we are a pure player of accessibility. And we based in Paris and we operating in Europe and U.S. and Canada today. Well, we're going to talk about PDF today, a few words about also DocAxess. We have, we are a team of 20 people today.

We are a tech company composed of engineers, PHDs, and and developers. So one of the main issue we're going to, we're going through today is well, PDF, well PDF remediation is a huge, huge issue on the web today. Indeed. I mean, if we if we look after the disability figures and how visual impairment affects a very large number of people in the world, we're talking about 16% of the world population concerned about that.

It's one in six people. It's a lot, lot, lot. And still, it's very a lot of people are unaware about that. And most of the editors are still issuing and publishing a lot of content content in PDF format, which is not accessible for visually impaired people, for instance. So if you yeah, the figures are there. They're tremendous. It's, it's a real concern.

Well, one of the, one of the main problems in our industry is, well, we have trillions of PDFs online. And addressing this issue is, is it's a mountain to overcome, really. So one of the main problems is how can we scale up with technology to address this issue? So this is the thing. This is where AI is helping to breakthrough and, and make some major enhancements in terms of efficiency and delivery.

We, we build up an outstanding and genuine technology based on three years of research and development, with the help of French labs and, and one of the best, I mean, engineers in math modeling and just cutting edge AI unlike any other on the market. They, I mean, we managed to basically modernize how the human eye is identifying the content on a page, whatever the layout of the PDF.

So when you when you look at a PDF, you statically can define and see whether it's there is an image or a header or whether the text is over two or three columns and in in which order you will

read the different content you see in a PDF. We managed to translate it into math and models enabling us today to, I mean, to, to gain significant, improvements in terms of PDF remediation process.

We have a strong and sound track record of 18 years of digital accessibility. We've, we, we have in-house guy, I mean experts and engineers. And we've been rolling out this technology since last year after three years of research and development for very large corporates that you may you may be familiar such as Hermes, Vinci, LVMH or Total.

So this... we we are a very... DocAcess is a very... I mean a very renowned player and pure player of the PDF remediation. We we are dedicated to this purpose and we have very sound experience in delivering a lot of very different documents, such as financial reports, annual reports, SEC filings, 20F. We have... so we've been outstanding also track record in documents, documentations, archives and also utilities PDF such as billings invoices, contracts and, and also payslips.

If you, if you look at it, payslip is a it's a major concern because a lot of people today can't even access to their personal info without asking somebody to read or explain to them what what is on the... what is on contract or payslip. We built in-house PDF checker based on our expertise and knowledge, and it's available online. I'm inviting you to, to, to test it and send us any feedback you may have.

So, so basically how's... how's the AI system we've developed. We, I mean, it's very straightforward. You, you will upload your document on the platform and will automatically detect and analyze your document according to reading all the titles, headings, page numbers, lists, etc., and etc.. We one of the... So how do, how do we solve this scaling issue?

One of the, one of the truth is the AI is a third party partner, to be honest, because it's doing 80 to 90% of the work. And of rest is finalized by one of the experts manually. It would be not... It would be false... and and it is and it would be a mistake to claim that AI can deliver 100% compliant and 100% usable PDF for screen readers for various reasons.

Technology is is mature, is still maturing. We, we're working very hard to, I mean, to progress and it's, it's an iterative work. But the final one, the final mile need to be, to be set by a, by an expert who will address texts, text out. It would provide contextual elements. He would also check and that AI had properly tagged the documents.

So this is a guarantee of quality assurance for, for you when using AI systems is if someone tell you that thanks to thanks to the AI system, they can deliver 100% compliance for document. So that's a red flag for you. I mean we, I mean, at DocAcess we using AI is, as I said, like a, like the, like a partner, who is part of the process.

But still we have it a hand on the wheel making sure that we, we deliver in the right the right file, and which is, consistent and relevant for, for the end user. One of the main, one of the main

advantage I would say, of our solution is it's a, we built a SAAS platform, so it's a very plug and play and smooth in terms of process.

You just have to drag and drop, upload your, your, your documents. The machine will, will ultimately automatically process your document and and you will have a... and you will be notified when the document is ready to, to download. So thanks to, to, to the platform you, you can, I mean you can also improve the governance in terms of PDF remediation within your organization because you can provide different ways for different people editing PDF.

You can speed and automate your your process in-house and we are providing multilingual customer support, in various languages. Obviously in French and English, but also in other languages, available on the platform. We've just, this technology's been, is been trusted by governments in France, in Europe, also in Germany, in Canada, and also for some clients in the US.

But before, but we need to also you, when you talk about PDF remediation, you need to look beyond technical performance. You need to be centered to, to the end user. One of the, one of the main goals that you need to be, that you need to be a very, a attentive is how your vendor is making sure that you're compliant with PDF/UA standards and WCAG. This is one part. But you need to also look at how it's going to make sure that your document is going to be very usable for your people. So you need to look after what kind of tests and techniques you use to, to, to make sure that what it is delivering is, to to you. And also what is the will testimony and promise in terms of artificial intelligence.

If he tells you that again, it's delivering 100% compliant by them, I think it's a go away and run fast and go away. We need to be very humble with the technology. Again, we, we claim that thanks to the, technology we developed we're making substantial progress and productivity efficiency in terms of delivery. For instance, we, we can deliver, I mean, a hundred thousand pages in a very, matter of days, or hour, for instance.

But it's not to, we still need to, to finalize, to read the final file manually. If you pass your PDF with accessibility checker, it doesn't mean that it's accessible every single time. So this is something that you you need also to be, to be wary about because like any automated tool, they can't, they can't really identify which, which is very, consistent and relevant, in terms of content and whether it would it has been tagged properly in a consistent way.

So those are quite a few few items to, to look after. We, we put on the, on the stage a couple of documents. I'm inviting you to, to look after those documents. I'm inviting you to also, contact us regarding our platform would be, I would be more than happy to discuss how we can help you in addressing PDF remediation.

Okay? I mean, we'll be happy to discuss also about, I mean, client case, how we streamline their processes in terms of remediation and how much they can they gain in efficiency and cost while working with us. One of the last points I wanted to to emphasize on is... Okay, being

trained on PDF remediation is one thing and making and addressing at scale is another is another part.

Well, you can, you can have a very good Word or PowerPoint or, or even InDesign file in terms of accessibility. And when exporting to into PDF, it doesn't make it accessible. You still have to finalize it with Acrobat Pro and you still have to test it, to be sure that what you've exported in PDF is okay in, in terms of your user experience for screen readers.

So when you consider this, you, you will... I mean, I'm sure that you you... It is very time consuming for you to, to make sure that all those steps and processes have been addressed. And one of the, I mean what we realized, for, for the past 18 years is that remediating PDF is a long process, is quite tedious, is very tedious and time consuming.

If it's, if it's not your core business, I am inviting you to contact an expert done that is doing this on a daily basis and are... and are well equipped in terms of tech and platforms. Please join me in the chat, which will be available. I'll be more than happy to add one subject to you question and see how we can help you to, to, I mean, to, to address this issue of PDF remediation. Thank you very much and looking forward to speak with you. Buh-bye.