

How to Fully Automate Accessibility in High Volume Document Processing

[00:00:00.39] So good afternoon, everybody. And thank you for coming to our webinar today on optimizing high volume PDF processes with automation. My name is David Herr. I'm the Vice President of Enterprise Solutions here at Allyant and I'm looking forward to presenting our information today.

[00:00:17.16] At the end of the webinar, I will be taking questions. So if you hold your questions to the end or if you put them in the chat box, we'll be able to then respond at the end, and go through and answer any questions that you have. So with that, let's jump into our presentation.

[00:00:34.50] So the agenda today, I've made the introduction as to who I am. We're going to then talk about the overview of the problem of volume digital documents. So we're kind of going to go into detail on that because it's kind of one of the last frontiers of accessibility. When you've done your static marketing documents and you've done your web content, you've done your videos for captioning, what do you do with those documents that are generated by various billing systems and systems of record?

[00:01:02.91] So after we talk about digital documents and how they're generated, we're going to then go into a screen reader example. I have an excellent example of using a screen reader to process an invoice-type, a statement-type document, and it really kind of shows what this really looks like in the real world to somebody using assistive technology.

[00:01:23.25] I then have some examples in both banking and government of how automation solutions can be built and developed to solve the problem. We'll talk briefly about some of the laws and regulations, alternative formats, and accessible documents. We're going to then talk about how this can be integrated into your existing workflows, because that is critical to being able to adopt an automated accessibility solution.

[00:01:46.55] It can't interrupt your current processes. And then finally, we'll go into the specific solutions that we have around automating remediation for volume documents and also some testing and alternative format discussions. So with that, we're going to jump into the conversation.

[00:02:04.16] So my overview of the problem. Your typical enterprise organization is generating tens of thousands, millions of documents. We've run into organizations that are generating hundreds million statements a year very easily, if you're a large health care or a large banking organization.

[00:02:23.39] These are normally generated from these customer communication systems, the OpenTexts of the world, the Quadient-type systems. And despite efforts in many of these systems to address accessibility, we find that they don't properly tag documents, despite the claims. Just because a document's tagged, or just because they throw the PDF/UA flag in the document that says the document is accessible doesn't mean that it is truly accessible.

[00:02:55.28] So tagging does not mean accessible. We hear that all the time. We're seeing the industry as a whole move from print to digital delivery.

[00:03:03.99] So the mass mailings of your statements, many times, if you-- with your insurance company or your bank, you can opt in for electronic delivery. So you don't have to have it come in the mail. It can be now an electronic delivery. But if you're going to make the document electronic, accessibility is now going to be a major concern.

[00:03:26.50] So how do you deal with those large volumes? How do you deal with hundreds of millions of documents? You can't possibly have a team of people to do manual remediation for millions of statements. So that's what we're going to talk about today.

[00:03:38.74] So these volume documents are typically generated from templated-based systems. Your typical customer communication management system has got a template that's been built. The template's been defined for the structure of the document. So it's laying out all the various characteristics of that document. It's then pulling the customer information.

[00:03:59.11] The name, the address, the account numbers, the balance statements. All that data is being pulled from databases and then populating that template. And then those templates are generated.

[00:04:10.04] And they can either be emailed, they can be put in a portal for electronic delivery, they can be delivered through various links, different ways they can be distributed. But these can be statements. These can be invoices. These can be explanation of benefits or other volume-type documents.

[00:04:27.46] And they need to be fully accessible, and not just tagged, but tagged properly. In many cases, we also see that not only do you need to have them in a PDF-accessible version, but you have to have some means to be able to convert them to Braille or large print for clients that have opted into alternative formats. And many states and organizations we're seeing are starting to require that you have a means for your customers to be able to ask for, or opt in for an alternative format for documents.

[00:05:03.27] So I'd like to use this sample of, basically, a bank statement, or a credit card statement. This is something our partner, Solimar Systems, has used this for one of their examples. And so I like to use it as a really good example of your typical document that's been created in a CCM-type system.

[00:05:23.21] So you've got this statement. You can see it's got dates in it. It's got account numbers in it. This one's been de-identified. It's got the credit balances and the previous balances. So it's your typical statement.

[00:05:36.19] It's got an address. It's got a barcode for the mailing. So this is what we're going to use for the sample with a screen reader.

[00:05:45.88] So what we've done is we took this document as generated from a typical system, and then we've taken the same document and tagged it properly for accessibility. So the first version is basically the document's been tagged, but not with any system that's actually going through and ensuring that the document is truly accessible.

[00:06:05.74] So the sample is going to show us and let us hear what the screen reader experience would be for somebody that would encounter the document that is not tagged correctly and then a properly tagged document. So with that, I'm going to start this and we're going to listen to this for a minute here.

[00:06:27.83] SOLibank Platinum Visa. Graphic three blue lines, S-O-L-i bank. Monthly statement. 123 Main Boulevard, San Diego, California, 92101. If you have any questions regarding this statement, please call linksolibank.solimarsystems.com.

[00:06:54.69] [AUTOMATED VOICE GLITCHING]

[00:06:56.68]

[00:06:58.98] Sean Connery. Credit limit. Available credit. 1007 West Street \$10,000, \$9,004.63. San Diego, California. 91903-4112. Current balance. \$995.37.

[00:07:25.31] Payment dates and balance summary. Payment total current. Previous interest payments/purchases/new due date, min payment, min payment. Balance May 28/21 amount plus credits charges plus balance. May 28th \$2,119.91. \$19.91. \$55.25. \$2.22. \$55.25. \$993.15. \$995.37.

[00:08:11.02] Heading level one, page one, SOLibank Platinum Visa monthly statement. If you have any questions regarding this statement, please call 1-800-555-1234/619-555-1235, TTY service 1-800-555-4321.

[00:08:38.19] Sean Connery, 1007 West street, San Diego, California. 91903-4112.

[00:08:50.66] Heading level two, payment dates and balance summary. Payment due date, May 28 '21. Total min payment, \$19.91. Current min payment, \$19.91. Previous balance, May 28/21, \$55.20.

[00:09:12.18] Interest amount plus \$2.22. Payments/credits, \$55.25. Purchases/charges, plus \$993.15. New balance, \$995.37.

[00:09:36.95] So as you can see and hear, the-- and this was using NVDA, the free screen reader tool that many people use that's available for download. So using that screen reader, you could really hear and see what that experience was like. If I was trying to navigate this statement and I was blind and I was using a screen reader, the untagged-- or the document that was tagged incorrectly-- as you can see, the reading order was all over the place. It would start reading the name, and then it would read a column on the right, and then it would go to the left, and then it would go to the right.

[00:10:10.16] Completely out of order. When it got to the point where it was talking about the payment and the balance summaries, you had-- it was reading across the screen. So you had no way to know that number that was being read was your current minimum balance that you, or the minimum that you owed, your payment due, those kinds of things.

[00:10:27.29] Completely gibberish. They tried to read the bar code that was used for mailing. Again, completely bad experience.

[00:10:34.61] But with a properly tagged document, I can navigate this and find the heading level for what I'm looking for. I want to-- need to find out what my minimum balance is. I can navigate to that quickly with the keyboard and then be able to read the screen, and listen to the screen, and go from there. So let me go to the next slide.

[00:10:52.95] So let's talk about laws and regulations now. There's been a lot of stuff going on lately with the Department of Justice and with enforcement of the ADA laws and the state and federal laws around digital documents and accessibility. As I said earlier, many states are now starting to see that you do have the right to request your statement from your utility or from your insurance company in Braille or large print, and that they can't charge you for that. It's free of charge. I know there's been a lot of activity in California around this space and organizations have been coming to us to be able to provide alternative formats because of the demand.

[00:11:33.96] And the Department of Justice has recently come out with a ruling in the municipalities and state and local government around ADA requirements. They've laid out the timeframes and what's exactly required to do to make sure that your web content is accessible. We're seeing the same thing in federal space. We're seeing the same thing in commercial.

[00:11:53.14] It's not just some of the lawsuits that have been going around that are driving it, but now it's the laws and the regulations are all starting to bring this to the front. And we're seeing a lot of activity and a lot of clients asking for these services. So let's talk about two examples of how automation can be built.

[00:12:10.73] This example's in government. Can't give you the names of the organizations, but I can give you some ideas of what was done. So we had a federal agency that had some volume-generated documents. The data was stored in a database and they were generating PDFs against the data in the database, and it was a high volume, to be able to do manual remediation.

[00:12:32.56] If I recall, this particular one was less than 100,000, but it was a pretty high volume of documents that they needed to find an automation solution. So we used one of our template-based tools where we actually built a template that would then generate the properly tagged PDF. So that template was able to go in, tag the heading levels correctly, tag the paragraphs correctly, make sure tables were defined with header rows, and header columns, and tooltips, and all those other things that need to be created within the document to generate a fully accessible, properly tagged document that would work well with a screen reader.

[00:13:11.83] And in fact, the way our tool works, it also generates compliance reports in addition to the actual tagged documents. So this worked really well in this government situation

to give them something to generate their documents and know that they were accessible and they met [INAUDIBLE] compliance requirements.

[00:13:29.32] In addition, there's a US major bank that works with Allyant. And one of the things that we find with working with data like banks and insurance companies is the requirement for really high levels of security. So we have two secure data centers in North America, one in Ottawa, one in New York, where we process all the documents that are sent from our clients that require a highly secure data center environment.

[00:13:58.70] So it's SOC2, it's PCI-compliant. We take in the files through a secure FTP process. They sometimes come in as AFPs, or they can come in as PDFs. But then we've built an automated tool that we call A Docs on the back end that will then, behind our firewall. It's not connected to the internet.

[00:14:19.30] We do the processing of these documents, generate Braille, generate large reflowed word documents. And then in many cases, we are then distributing those communications to their clients directly for them. So we're mailing out the Braille statements, or we're mailing out the large print.

[00:14:39.02] We're doing about 9,000 communications a week for this particular client and about 75% of it being large print. So we're seeing more and more people that have low vision, not full blindness, that are asking for alternative formats such as large print.

[00:14:55.57] So one of the things that I wanted to talk about with these types of systems is the need for this to fit into your existing workflow. Here at Allyant, we've been involved in automation solutions and been in the PDF accessibility space for well over 20 years now. And early on, a lot of our technologies were built around creating the template. So we built some software solutions where we create the template for your accessible document that you need to generate.

[00:15:26.99] But that doesn't work necessarily in the real world. Many of our clients are using a customer communication management system, a CCM system. They're using OpenText. They're using Quadient. They're using one of the other similar tools that's deeply embedded in their organization. It's deeply embedded into their workflow.

[00:15:46.03] The customer communications have-- marketing is touching them, finance is touching them. Other departments need to be able, in the approval process, to approve final documents before they are generated and then distributed to clients. So because of that, they have to be built in these systems and work within these systems so that the approvals can take place.

[00:16:11.10] So we can't do the template. It actually has to be almost after the fact. So the customer communication management system is doing the generation, creating the documents. The final document, before it's then distributed to the client, is what then needs to be tagged. So it's going to work in parallel to an existing system.

[00:16:29.10] Sometimes your files may need to be processed on the fly, or maybe they'll be done in a batch or on demand. We've seen many clients where they're putting the files into an archive that the customer retrieves them. They can retrieve them then in an accessible format using our technology. And then you need to be able to monitor that the output from the system is meeting your requirements from an accessibility standpoint. It's some kind of a QA process.

[00:16:56.98] So this drawing kind of shows how our systems work and some of the various things that can occur. You've created the final PDF document. So on the left-hand side, your CCM system has now generated a document. It would now either eventually go to be printed. It would either go to a portal and for e-delivery, or it might be archived so you have copies of it that somebody could retrieve in the future.

[00:17:24.22] So before the file is distributed to those systems on the right, it's going to run through our system in the way that we build this. So through an API call, the file is going to come in, it's going to be tracked, it's going to be monitored.

[00:17:39.82] We use dashboarding software from our partner Solimar Systems to be able to track the workflow of the documents. And then the documents are delivered to our platform. Our platform can do testing and validation. It could convert to Braille, it could do large reflowed print, or, of course, accessible PDF.

[00:17:58.52] In this case, we're talking about accessible PDFs. So the accessible PDFs are going to be delivered on the other side out of the system. Again, this could be done on demand or it can be done in a large batch. So you could produce all your statements and then they can be run in batch mode, made accessible, delivered to a hot folder that then is retrieved by your systems and then delivered into your workflow.

[00:18:20.24] So it has to be parallel. It has to work with your systems. It can't-- it's not going to replace your system. This is going to supplement your systems.

[00:18:31.18] Security is one of the biggest concerns in these automated solutions. So in our case, the way we've done most of this, if we're doing the processing of your documents like we do for those major banks I talked about, that is behind our firewall. It's in our secure data center.

[00:18:49.58] We're going to take in your output documents. We would do the processing and we would do the delivery. But in many cases, our clients are going to be doing the hosting.

[00:18:58.89] So you would put this solution behind your firewall. You can run our software tools and access the APIs behind your firewall. So all the security and all the data is kept safe and secure in your own IT systems.

[00:19:12.77] Again, since that format may not be PDF, it could be coming out of a mainframe in AFP format or something similar to that. So one thing I strongly recommend is that you look at-- if you are having another vendor post your solution, you got to make sure that they're certified as secure.

[00:19:32.46] Do they have HIPAA compliance? Are they SOC2 compliant? Are they PCI/DSS certifications? We have that in all of our data centers. We're trusted by banks, and hospitals, and health care systems, and the federal government in that regard. So you got to make sure that you've got the security if your vendor is hosting your data, or you put this behind your firewall to ensure that it's secure.

[00:19:59.97] So let's talk about what can be automated. What can't be-- I'll talk about what can't be automated before I talk about what can be automated. What can't be automated today is static documents, so one-offs. So your marketing brochure that you've created an InDesign and it's going to be stuck up on your website. It needs to be accessible. It's quantity one.

[00:20:23.82] So today, the automation tools, it's not going to make sense to build a solution to remediate a single document. So right now, we don't recommend any automation solutions for low volumes. But if you've got higher volumes, these are the things that can be automated.

[00:20:40.78] So accessibility compliance testing. You can test the documents for accessibility. They can be scanned, they can be verified, they can go against WCAG or PDF/UA compliance checkpoints and be tested, and you'll get a accessibility report and know what your next steps are.

[00:21:00.60] PDF tagging can be automated. So if you're using a templated-based solution that's generating high volumes of PDF documents, that's ideal for automation of PDF tagging to make sure they're properly accessible. Braille can be automated. Again, if you just need a single Braille copy of a textbook, you're not going to automate that.

[00:21:23.82] But if you're doing Braille statements, absolutely. Explanations of benefits and things like that, those can all be created as Braille and be automated. And then large reflowed print.

[00:21:36.46] And many people will misunderstand that and think that, well, you're just blowing up the font, or you're just basically expanding the document, size-wise. Well, you can do that, and some organizations do do that. But a better experience for your clients is to actually reflow the document. So a three page statement might be 10 or 12 pages if it's been reflowed.

[00:22:02.06] But the columns are going to be moved, the font is going to be, yes, increased in size, but you're going to be able to read through the document. Each column might be on a page, instead of it all trying to be blown up and trying to be read. So it works much better and we've been working on this for many years. I'll talk more about that in a minute.

[00:22:19.83] So let's talk about how you can do the testing first. We have a tool called CommonLook Clarity that's used for testing of PDFs. So this is our automation tool for the testing of PDFs.

[00:22:32.62] This can be used as an API and test files one off. You can actually build a system where if your system generates a PDF and you need to test it for accessibility, the tool can test it. If the file is accessible, it can then be delivered into your workflow. But if a file is not accessible

and it is not properly tagged, it can then be sent to one of the automated tools to add the accessibility tagging.

[00:22:58.47] We can also use this tool on a bulk of documents. So let's say you had an archive of old documents and you wanted to test those for accessibility. You can run this tool against that block of documents. You're going to then get compliance reports to WCAG, PDF/UA and other standards to ensure the documents are accessible. And then you can just make your determinations or your next steps after that.

[00:23:27.42] Yeah. So here's an example of a report. As you can see, we've tested this government client. You can see the URLs for where the files that were discovered.

[00:23:37.08] It tells us if they're tagged. It tells us how many pages the document is. It tells us if there's any PDF issues with the structure of the PDF.

[00:23:46.95] Is it a corrupt file? Is it a file that's a clean PDF document? Is it file protected so you can't write to it, which is-- you can't tag the document if we can't write to the document.

[00:23:56.35] Tells us where the file was created, what platform was it created on. What standard did we test against? In this case, I tested against WCAG 2.1.

[00:24:05.92] And then we have a verification of the status of these failed. We can then get detailed reports on each individual file, which will tell us where it failed. Was it an alt text issue?

[00:24:16.19] Was it reading order? Was it heading levels aren't defined, things like that. So it'll give us the details on where the accessibility issues exist.

[00:24:28.68] So now let's talk about PDF remediation. And Allyant uses our CommonLook brand of products around PDF accessibility. So everybody's using the AI term now and it's pretty exciting to see what AI is bringing to the real world. And AI definitely has a place in accessibility.

[00:24:51.24] We're seeing it more and more every day in new tools, and solutions to help people with disabilities and to ensure content is accessible. So in the document space, CommonLook's worked on a number of different products around PDF accessibility and automation. So we have a product. We call it CommonLook AI where we take the documents, we take the samples of the documents generated from your templates.

[00:25:19.67] So let's say you're using OpenText and it's generating your statements. So those statements-- we will get samples of those statements, and we actually tag them properly for accessibility in a manual process. So we take your samples, we do the manual tagging. We test them and ensure that the results are accessible.

[00:25:38.71] We try to get variations of how the documents might come out in the real world. So let's say the statement is normally two or three pages, but it can be as many as 10 pages. We'll

get samples with 10 pages and seven pages and five pages, just to see what structure changes might occur to the document in the real world when they're generated from your systems.

[00:26:01.23] Once we have those samples and they've been tagged, we then actually build an AI model. We test and build that AI model, and teach it, using these remediated documents, how to do the tagging and what needs to be tagged. So we eventually build this AI model. We can now run against your volume documents through an API call.

[00:26:21.58] So these documents can now be tagged on the fly, real time, either on demand in batch or as needed, where the tagging will be added to the documents. It generates an accessibility report and the tagged document. And then these documents can continue through your workflow.

[00:26:40.48] We also have a second tool which is similar to CommonLook AI that uses AI for the analysis of the document. So you would open up that sample document that I talked about in this tool. Once you're in that tool, the AI product's going to go through and identify the heading levels, and the tables, and the paragraphs.

[00:27:04.05] And the rest of the time, it'll ask you for things that it can't do such as, is there alt text, or is this just a decorative-- I show an example on the slide. The man shaking the hand on the right, is that just for decoration or is that conveying information? If it's conveying information, you would type in the alt text for that image.

[00:27:22.53] Once you've confirmed that the tool has properly identified the content, you've made any changes that need to be made, this, then, becomes the automation template that we can then run at high volumes, your documents against, again, through an API call. So unlike the other product where we actually teach it and build an AI model, in this case, there's some manual intervention.

[00:27:45.05] It's very useful when there's a lot of templates that need to be created and there's a lot of work that's going to need to be done to generate it. So we can basically hand this over to your development team to take and build templates for multiple documents. So we have solutions for high volumes and we also have solutions for high numbers of templates where it's not going to be realistic to build 8,000 models.

[00:28:10.23] And where the technology is going is we're moving more and more towards more generic, more broader AI models that will be able to handle a multitude of documents and not just be optimized for a particular type of document.

[00:28:26.24] So in looking at our automation solutions, they are designed for the enterprise. They are designed to be API-driven. So the integration in your existing systems is take the file, run the API, send us the file, the tool runs, processes it. It meets compliance for Section 508, WCAG 2.1, 2.2, 2.0, PDF/UA, HHS standards.

[00:28:49.97] It is built on the latest technology. So we are using machine learning. We are using true AI through deep neural learning AI, thus the AI models that we build. And also, we have

template-based solutions. So we've really kind of built our technologies around what's the best fit for the client need, not we have this technology and that's the only thing we have. So we have multiple technologies in this space. We've been doing this, again, for almost 22 or 23 years at this point.

[00:29:20.32] We also then have the dashboarding to be able to track the workflow throughout the system so that you can get reports. You can get access to-- what's the-- as a QA process, as we generate the accessibility reports, how's that going? Have we made some changes to the documents that are involved? We're going to have to go back and make some changes to the model, and that can all be tracked through the systems.

[00:29:44.02] So let's talk about accessible print remediation and production. So as we said before, there's a print side to accessibility that goes beyond the electronic version, which we typically deal with on the PDF side. So Braille has been around, and there's various versions of Braille.

[00:30:06.22] And I think we're seeing that more and more in the modern world, modern education space, they may be moving away from Braille. But there's still a need for Braille and people that request a Braille version of their statement, or their insurance policy or things like that are going to request it and you're going to need to have a way to provide that to them.

[00:30:28.64] We've been doing this, again, through our T-based subsidiary for many, many years. So we're very involved in the various versions of Braille. We can support them all.

[00:30:37.88] And on the reflowed large print side, there's never been an official standard for reflowed large print. So what we've developed, in conjunction with our customers, is what's the best way to create large print that is a good user experience for their clients? Again, just taking a document and blowing it up to a larger size sort of works.

[00:31:03.87] But you exceed the size of normal paper and things like that. You're not going to send them a poster of their document. So to be able to take the document and reflow it into what fits on 8.5 by 11, but now have the font size increased so that it's an easier reading experience, is a good usability best practice, and it's something we've been following for many years.

[00:31:28.87] Yeah. And then finally, how do you prove compliance? One of the things that we've built into all of our solutions where you can test for compliance on the PDF side is our compliance reports. CommonLook generated these many years ago.

[00:31:46.93] And there's a couple tools out there today. And not picking on Adobe and not picking on others, but passing some of the previous accessibility tests for documents has always proven to be rather difficult. You can pass the Adobe Accessibility Checker. You can pass the Microsoft Word accessibility checker and have a fairly reasonable, accessible Word document.

[00:32:10.22] But then when you save it as a PDF, if it lost the tagging or it didn't tag it properly, you now have gone back to an inaccessible document, and that defeats the purpose. Passing the Adobe checker may mean that some of the testing that they've done, you've completed and it's

passed, but it can still have all kinds of accessibility issues that are going to be a bad experience for your customers.

[00:32:33.38] So on the CommonLook side, we created our accessibility reports to go through the standards checkpoint by checkpoint and confirm that the documents are meeting the requirements for accessibility. We guarantee it. We guarantee compliance. If you pass our test, you're going to be able to, with assurance, know that your documents are accessible.

[00:32:57.78] Even with our AI tools, because the models that we built are based on manual remediation where the images that were in the document-- I mean, on a templated-based document, that logo doesn't change. It's the same logo. So the alt text of that logo has already been manually entered in, confirmed by a human that it is accessible, and now that's just being applied. When that logo is on that location, on that page, it adds the alt text for that logo. So with our compliance reports, you can be assured that you have full accessibility.

[00:33:31.29] So I want to thank everybody for attending today and I'll be open to questions now. We can pull what's in the chat box or we can take a look at what comes in while we're-- our remaining time here. So I look forward to your questions. Thank you.

[00:33:44.16] So as we've been going along, I've been answering questions along the way. Got some additional questions. There was one question about what is the screen reader that we recommend that you use. We do most of our testing using NVDA, primarily because it follows the standards.

[00:34:03.39] I know a lot of people also use JAWS, and there's nothing wrong with JAWS. JAWS does have some different behaviors that don't necessarily follow the standards, and hopefully, at some point, they will follow the standards correctly, because you sometimes get into a situation where people will want to tag a file to work better with JAWS. But if you tag it to WCAG or PDF/UA standards, then it should work with any assistive technology. So we do recommend NVDA.

[00:34:33.69] I have another question was asked about where is the report housed for each document? The way the AI tool works is we'll typically take in the files through an API call. They can be dropped in a hot folder or something similar.

[00:34:49.78] We would then process the files and then the files would be returned to a hot folder. So you would get the tagged PDF and then a matching compliance report for each file. Obviously, using document management systems, you can move the reports somewhere else and process the PDF documents to deliver them to wherever their final destination is.

[00:35:14.82] I have a question that says, do you have document templates for meetings, agendas, and minutes? And what is available for engineering documents? Actually, it really does not matter, as far as what the document template is.

[00:35:29.35] As long as it is a templated type of document, it's a good candidate for automation. So let's say you had meeting agendas and minutes to the meetings. If you can follow

the same format, the data can be different, the length of the tables can be different, all of that. But you're not changing the structure of the document and it's a templated style document, we can build an AI model around that particular document.

[00:35:55.96] When it comes to engineering documents, that gets a little more complicated, primarily because if you've got lots of engineering drawings that are going to change, and each of them will require a subject matter expert alt text description for the document. So you may have a-- let's say you have a model you could build for a structured engineering document, but if it's a different construction project that's going to have different engineering drawings that are involved.

[00:36:28.80] So you're not going to necessarily be able to use AI for something like that. You may have to go in and type the alt text for those things. Or what we've seen in many cases is if it's just not feasible-- I mean, lots of times an engineering document will have all the descriptions in the text of the document and then it'll include a drawing for additional information.

[00:36:49.86] It may be too much to convey in an alt text drawing. I mean, it's just impossible in some cases to describe every minute detail of that image. So you may have the basic parameters as to what that is describing and then include contact information to get more information, if it's necessary. So you want to make sure that that's something that's feasible. So engineering drawings may not be the best use for automation, but definitely your meeting minutes and definitely agendas and things like that.

[00:37:22.54] Let's see what other questions we have. Can we Use CommonLook to output information from the remediated PDF into a Word document? Not sure exactly what you're looking to do there.

[00:37:40.98] I mean, if you created a PDF using CommonLook and then you merged it back into a Word document, it's not going to bring over the accessibility tagging that's in the PDF. So I'd be happy to answer that question offsite-- or offline here and we can get some of our trainers to talk to you and find out exactly what you're looking to do.

[00:38:05.03] Somebody asked about, does our 100% guarantee compliance seal protect clients from liability from a legal standpoint? The guarantee that we do is around that the document meets the compliance standards. The documents are going to be 100% compliant with WCAG 2.2, 2.1, PDF/UA, HHS standards and we guarantee that.

[00:38:32.46] Over the years, have we had to support our clients and prove compliance? We have, and we were there to help them with that. So as far as any challenges to the legal standing, as far as the documents being accessible, we do stand behind the work that we do.

[00:38:49.28] Question on can I use shortcuts in CommonLook create workflow with a stream deck. There's a number of shortcuts that are built into CommonLook, and I know that some people do do some automation or scripting of how CommonLook PDF can be used. It's separate from our AI products. I'd be happy to have you talk to one of our trainers offline that can answer that in much more detail.

[00:39:16.50] There's a question about on-prem solutions. Basically, the way we do that is due to security, we have not been selling our automation tools as a cloud-based service. Most of our clients, they're dealing with financial information and they're dealing with personal information and things like that that they want to keep behind their firewall and keep secure.

[00:39:43.47] So we do offer the product as an on-prem solution. You would put it behind your firewall. You would install it with all of your security parameters behind the scenes and you can then do API calls against the solution to generate your accessible documents. And then nothing leaves your firewall, nothing leaves your data center, from a security standpoint.

[00:40:05.85] The way it's priced and costed is really based on a number of parameters. How many models are we building? What's the annual volume of pages that you see generating by the system? And based on that, then we work up a price that covers the model, covers the QA, covers building the models, and then supporting the tool in production. So that's something we can certainly talk to you, Philip, offline, if you'd like.

[00:40:32.31] I have another question that says we need financial documents, municipal budgets remediated. Is there a program that can process this? Absolutely. We do a lot of manual remediation here at Allyant for those one-off documents.

[00:40:46.57] So you've got that 800-page municipal budget that you do once a year and it's built out of Excel, or it's built out of your accounting system, and it can be a very complex-- lots of tables, very complex for somebody to manually remediate. But we have solutions to be able to remediate that document. We can do that at a price per page and it can be something you can easily budget to do, and then takes that worry out that you've got to figure out how to manually tag a one-off document every year that can be very complex and difficult to do, if you're not highly skilled at document remediation. So I'm happy to talk to you, Peter, offline. We can work out what that would cost and give you some details on that.

[00:41:27.77] Question that says, I have a couple of document formats. I'd like to see if they'd be good candidates for automation follow-up to a range of tests. Absolutely, anonymous.

[00:41:37.92] We'll need to get your contact information and then we'd be happy to be able to show you how the product works, look at your documents, tell you if they're good candidates, and give you all the details. So reach out to DHerr@Allyant.com or to our sales team here at Allyant and we'd be happy to help you there.

[00:42:01.13] Another question. It says, in a self-hosted model, can the customers develop templates, or does it always have to go through Allyant Professional Services? Well, the interesting thing is we don't have a "one size fits all" technology. And this technology is constantly evolving.

[00:42:19.67] So the way our AI product works, we do build those models, because it does involve some pretty high-end computer back-end systems to generate those models to teach the system after we've done the tagging. If you were tagging the documents, and then you were building the models, and then they had accessibility issues because things weren't correct, we

can't stand behind that for the AI product. But we do have another product that also is an automation solution that works very similar to what the output is from CommonLook AI, that the customers can work on the models and can do some of that development.

[00:42:59.22] So happy to talk to you about that offline. Share your contact information. Reach out to the email addresses and we'd be happy to walk you through what that means and show you how that would work.

[00:43:11.04] Another question that says, is the content used in the language model for AI also secure? Yes. The AI model is developed-- we would develop it using test samples. Test samples can be de-identified, so they don't need to have-- they won't have any private information in those models as we develop them.

[00:43:33.08] Then, when they are hosted by you, you would put them into your production environment. You would then feed it live data that has, you know, private information in it and you'd get your output. But the system's not storing that and you're actually hosting it in your data center. So you don't have to worry about security.

[00:43:49.84] We're not containing any information in any kind of a public AI system that you would be sharing. This is all self-contained. This is not reaching out. It's not using OpenAI or any other tools like that.

[00:44:04.44] Another question. It says, we are a government and we have thousands of documents. However, many of them are one-off documents. Would they still be used for automation?

[00:44:14.07] Unfortunately, not today. I think the Holy Grail of automation is that we'll be able to build a single model that any document style can be run against, and when that type of technology is actually available, that's going to be a really great thing from the standpoint that you could run a file against the tool and you're going to get back a tagged document, irregardless of the layout format, style, structure of that document.

[00:44:40.36] But right now, today, we do have to build models based on similar structure. So we can build a letter template and take in and teach the tool the various structured changes in a letter, and then get really high quality output based on that model. But if I sent a budget file against that, or if I sent a bank statement-style document against that letter model, it's not going to do a good job of tagging it. So your one-off documents today would not be a good candidate for automation, but that is something that I can assure you we are working on diligently to be able to build, and at some point, hope to offer that to somebody, I hope, to customers in the future.

[00:45:26.05] Can you handle power BI reports? And the answer to that is if they are template-based. So if the structure to those reports is repeatable and it's just the data that's changing, but you are not taking tables and adding 12 more columns in the next report you send over, yes, the tool can be used for structured template-style reports. And it really doesn't matter what the source is that they come from.

[00:45:57.24] A question on what-- do one-offs have to be templates as well to work well with the AI solutions? They don't. I mean, we can do a one-off document. We're not going to use automation, we're going to manually remediate it, because we'd have to manually remediate it anyways to teach the tool to build the model. So you're just going to manually remediate any kind of a one-off document. But it doesn't have to be template-based. It can be anything.

[00:46:23.49] Another question. AI solution helps with remediating PDFs. How does this differ from the remediation service? The difference is manual remediation services are you send us documents, we remediate them, we send them back with an accessibility report.

[00:46:40.71] The AI solution is for those PDFs that aren't stuck on your website and aren't public information, but are generated in high volumes thousands, tens of thousands, sometimes millions of these documents. So these are like-- let's say I'm the cable TV company and I'm generating your monthly statement, and that's available as a PDF. I'm generating millions of those. We're not generating hundreds of those.

[00:47:09.18] So those documents are great candidates for building an AI model and then being able to generate accessible PDFs out of those systems and then put those documents out in a portal for people to download, or to get their statements. So that's the, really, difference between remediating millions of PDFs or hundreds of PDFs using an automation tool versus manually remediating. Your typical utility is not going to want to send millions of documents to an outside service to remediate their statements.

[00:47:41.20] So this is something that can be done on the fly. It can be done in batch or it can be done on demand to add the accessibility, if somebody requests a file. So there's a number of different ways to put that into production and make it realistic and feasible. And it's a great alternative to manual remediation when it's a good fit.

[00:48:02.19] Somebody asked a question about how does your solution compare to the Adobe PDF accessibility auto tag API? Are they similar? Adobe has had that for a long, long time and they've continued to advance it, and add to it.

[00:48:21.50] But if you go into-- take PDFs and you run the Acrobat add tags, and you go into that Acrobat and then you look at the tag structure and you see all the errors that are in it, as far as reading order, as far as properly identifying content and tagging it correctly so that they match and they are correct and it's a good, usable, accessible document.

[00:48:48.08] From what I've seen with some of the automation tools that are out there from Adobe and from others, they still require some manual manipulation of the output. There's still errors and issues with how the documents are tagged, and there's things that have to be done to fix it. With our tool, we're going to remediate your samples and we're going to teach the tool how to tag the documents.

[00:49:13.79] It's not trying to guess what an H1 is. It's been told when the document has this structure in this area on the page, this is going to be an H1, and it's always going to tag it as an H1. It's not going to decide, well, maybe it's an H2, or maybe it's an H3, and we don't need an

H1. I mean, you get all those kinds of results with some of the automated tools that are just using machine learning and guessing at the content versus we are specifically teaching the tool how to tag your documents correctly.

[00:49:43.62] That's the other reason why we generate an accessibility report when we're done that will show us that there is an H1, heading order's correct, the reading order is correct, the content's correct. All text descriptions aren't just this is a picture JPEG kind of thing, but it's an actual description, things like that. Any other questions? Lots of good questions today.

[00:50:07.64] OK, another one. We've used the service side for hundreds of PDFs where we upload to a portal and receive a zip back. Do you offer an API of sorts to potentially automate some of this rather than manually uploading? So this is to do manual remediation.

[00:50:24.59] We have not done an API yet to upload to our portal, but I'd be happy to talk to you about that. And I can certainly see in dealing with-- I mean, many times, we're finding that people have files stored in Google, or they have them stored in other services. So to manually have to pull each file out and then upload it to our portal, I could see where that could be something that would be of use. So yeah, please reach out. I'd be happy to talk to you about doing some kind of an API to allow you to automate the uploading of files to the portal.

[00:51:00.03] All right. Unless somebody's typing some last answer here, I think that'll pretty much wrap it up for today. I really appreciate everybody attending. And we will get this slide deck out to everyone, because I know there were some issues with bandwidth and the quality of the screen.

[00:51:16.06] So we will get that out to everybody so they can review it. Please reach out to [Allyant Info@Allyant.com](mailto:Info@Allyant.com) or DHerr@Allyant.com and we'll be happy to talk more in detail and answer your questions further. And we look forward to speaking soon. Thank you. Take care.