

Keep, kill, combine: fixing intranet sprawl before AI makes it worse

Jarrold Williams, St James's Place

Cai Kjaer

So Jarrod, I got hooked by that title, Keep, Kill, Combine. I just love that slogan, so take it away.

Jarrold, Head of Digital Channels, St James's Place

Yeah, cheers Cai, thank you. So I'll just do the quick share of the screen and bring up and here, so hopefully you can see the screen now. So yeah, thank you for having me, an opportunity to kind of go through this presentation. So as you say, Keep, Kill, Combine and we were having a chat with Cai a few weeks ago and you said, I'd mentioned that phrase and how we're tackling a challenge here at SJP. So that is my title.

So my name is Jarrod, I'm Head of Digital Channels at St James's Place. I'm going to be doing a walkthrough of a very real live challenge that we're actively tackling. So how our three-year-old SharePoint intranet has already begun to develop the content sprawl and governance issues. And I want to share kind of what we're doing right now, kind of actively day to day to try and fix that and how getting the fundamentals right is critical before you can even consider using something like AI. And we often hear challenges around SharePoint as a platform and non-SharePoint users will moan and begrudge it.

And yes, it has its foibles, but intranets don't often fail because of the platform, they fail because of what sits on top of it, whether that be the content, whether that be the structure that you have, the governance that you put in place, even the question that was asked earlier, the purpose, why even have an intranet and not understanding that completely. And particularly in an era where we have kind of AI digital workplaces, that cost of kind of messy content is going to be kind of more substantial. So very quickly, who is SJP to set the scene? St James's Place is a UK based financial wealth management firm for financial services with FTSE 100. We have a large footprint across the UK as well as the Middle East and Asia.

And our workforce comprises of roughly 2,500 employees and then 15,000 of what we call the partnership. So these include partners of SJP, financial advisors and their support teams. And they're across 2,800 independent businesses from us. So they sit under the SJP umbrella, but they are independent businesses. And our SharePoint intranet is unique to a lot of people, maybe not to anyone here that has a more franchisee model, but we have to serve both of those audiences, both the internal employees and those externally facing

partners as well, the advisors.

And so that isn't, that means it isn't just a internal communications platform. It's actually the core primary information hub for those advisors in a very regulated industry. And as I say, across multiple international geographies. And that means that it's quite high stake for us. We need timely, accurate information.

We need up to date product information policies so that they can do their jobs. And they absolutely have to trust that what we're putting up is accurate. And I will say, and I'll mention this a few times, we have a federated publishing model. We have, like many of you might have, not a large central team, but we have lots of people around the business who are contributing to our content. That's great in terms of local ownership and expertise, but it does present us challenges around consistency, quality, and the currency of our content.

So as I said, before I really jump in, this is a live case study. It is not a, we've done this, we've got the award, I've got a trophy and everyone thinks we're brilliant. This is a snapshot of literally what we're doing day-to-day right now to try and fix it. I actually felt like I should have come maybe before the previous speakers because our process is kind of the start of that journey and how we're tackling the content sprawl and building some of the things that you saw earlier. So our intranet in numbers.

So to understand why this problem is really urgent for us, it's good to see how big and busy our intranet is. So we have roughly 45,000 pages and documents. Now to some of you that might go, well, you know, we've got very similar. We only had 15,000 when we migrated three years ago. So we've seen this exponential growth in the pages and content for a relatively young platform.

And that's across roughly 250 SharePoint sites that we say make up our intranet. So we get about 1.3 million page views every month because of the kind of partnership using it regularly. And we have around 500 content editors. Now again, that, you know, compared to 2000 previously, that seems smaller, but that makes up roughly one in five of our employee base is in some way contributing to content. So the vast majority are in our intranet editor community ultimately.

And we find that an average 23 minutes per session, that could mean that everyone is deeply engaged with all of our great content and they love it and they want to spend all their time there. Or it could mean they're spending a lot of time trying to hunt for things down and find something they can't find.

And there's a few things there, it probably means less to you, but our top sites just to show giving advice, advice, framework, investments, the reason I call that out is there are big critical sites for our partnership community because they're about the financial advice, they're about our products, they're about the investments that we offer. And you can see the kind of things that they search for, like fund centre, income sustainability calculator, very financial things that I have no idea what they mean generally day to day. So yeah, this is ultimately a massive kind of digital ecosystem.

And then we have other platforms as well, like our L&D system as well. And if you let all this go unchecked, if you let it all grow organically, you very quickly get a jungle. So why intranet sprawl happens? How did we get here so fast with an intranet that's relatively new? And I'm a fan of photorealistic Copilot generations of people battling SharePoint.

You've already seen one, here's another. So as I said, we moved off a legacy platform only about three years ago. We had a deadline to do that. And again, going back to the previous thing, we ended up having to do the whole content needs to come over as is in some places. That meant we had to carry across some technical debt, some content debt, but that was a trade off for us to launch on time.

So we were aware of it, but it kind of set the stage for where we are now. What I will say is we did extensive user research. So when I joined the business, we're already into the kind of migration process and we did lots of user research. And if I implore you to do anything, it is to do user research, use UX designers, use user researchers and develop something that is led by user need. So for us, that means that we understand that our fundamentals are right.

It's maybe just some of the kind of in the weeds where it started to slip. And because we had that federated model and perhaps we did like governance, we have started to see duplication and inconsistency because we have 500 editors, so 500 versions of good. Now we do have some, we have an editor community, we have rules and guidance. We even do the training as well. But we approached it through the way I explained it is very much a carrot in the carrot and stick analogy.

And what we realized afterwards is we shouldn't be using the carrot. We shouldn't be using a stick. We should be using a large back with nails in it. And that's probably the only way that we're going to make sure people stay on our governance. We have unclear ownership.

We have just been through a very large restructuring program. So we've lost members of our workforce. And as time has passed and those kind of things have happened, people have left. We found that ownership has become a bit more unknown. And frankly, if you can't name an owner, you can't trust the content.

And then we've also seen that we have kind of, we have a life cycle process. We have review cycles and approvals, but we're finding that people, because again, carrot and stick, we didn't mandate the archiving at the end. So the review kind of flags to people, but doesn't automatically remove. So again, that life cycle process isn't quite there. And the symptoms of that sprawl, I've already mentioned duplication.

We found in some page libraries, five or six versions of the same document, each team believing that it's theirs. Everyone convinced it's the right version. We found outdated content still live. And a great example of this is we found a piece of content that was a product that we hadn't offered for basically five years. So straight away, it'd been migrated and shouldn't have been.

But the reason it popped up is because it was being edited actively. And a well-meaning content editor had saw some typos and were going in and they were tweaking those typos and going, I'll just update the little blurb at the bottom because that's not right. Not realizing that by updating it, they're signifying to SharePoint that that content is still relevant and should be exposed at the top of search results. So we had that problem with outdated content that was actually being edited. We saw obviously search, and this actually searches the number one driver of our project is better search results.

So this is critical as kind of digital workplace specialists, internal comms, intranet managers. We tend to blame search when people can't find things, but it's rarely search that's the problem. It's usually the symptom. So search was doing its job. It's just the amount of clutter underneath.

And navigation, I mentioned, we did user research and navigation we think was solid, but down in the weeds on some of our sites, navigation was becoming bloated as more and more people tried to shove their important things towards the top. And as this happened, people started making workarounds, saving local copies, emailing people for documents rather than going to finding them. And we started to see risk creeping in. For us, that's really important. We're a regulated business.

If someone makes a mistake giving advice to a client, not only could it adversely impact that client, but it could have regulatory repercussions on us. So the massive business risk of this. Sorry, clicking the wrong way. So our three pillars for fixing this. We've got a project, it's part of what we're calling our tactical admin program.

And this is the kind of intranet cleanup project. So we're going three different pillars, a content audit, a governance and technical stream. So content audit, we're going to systematically review and restructure the content across our whole intranet, every single one deep breath of those 45,000 documents and pages. And that's where Keep, Kill, Combine comes in. We've got a governance work stream.

And probably for me personally, I am a disorganized mess often. So governance is not my forte. Luckily, I've got a great team around me. But we're going to have a governance work stream that's really important. And we're going to put in place the rules.

We're going to place clear roles. Every piece of content will have a responsible owner. It will have a defined purpose. There will be a lifecycle with creation, review, retirement, to ensure that we don't have that spoil again. And then we're going to have the technical fixes, things like tweaking search, adding extra tools that prevent this from ever happening again.

And we're doing it in a stabilized, redesigned, future-proof kind of three-step process, a three-step plan. Stabilize, stop the bleeding, quick fixes, easy cleanups to kind of reduce the worst clutter. Redesign to make some of those deeper changes on the structure, the information architecture, page templates, and then future-proof, ensuring that our model is more sustainable. So in 2029, I am not on a sloop festival saying, here's our project to get rid of 100,000 pages. So big picture here, we're going to tackle these issues, three core

pillars of work, lots of different angles.

Title of the presentation, Keep, Kill, Combine. So the first pillar of our content audit where the heavy lifting will happen. So we are inventorying, is that a word? Inventorying everything. So making an inventory of absolutely every piece of content.

And that's a bit of manual work and it's a bit of using data analysis and tools. So we pulled together a giant list of everything, including things like last updated, last views, the known owners, if they have them, and this is our core evidence base. And then we have this Kill, Keep, Combine framework. So kill, if the document is obsolete, it is unused for over a year. If it is duplicated elsewhere, or just doesn't have an owner, we will archive kind of without question.

Not going to go and check with people, we're going to archive those documents. Archive for a safety net, not necessarily delete. Keep, if it's current, if it sees strong usage, if it's business critical, and as a confirmed owner, then we'll keep it. One flag on that, we'll still do cleanup of that, some metadata and formatting. And also we acknowledge that sometimes, I put there high usage, we might have something that has very low usage, but has a regulatory reason for existing, i.e., we have to have it somewhere.

No one reads it, but if the FCA were to ever come and audit us, we need to show that we've put it on our intranet. So we'll also need to identify those things. And then Combine is just where we have the overlapping or duplicate content, we'll bring it together and try and reduce the amount of repetition that we have. Now, because we've got tens of thousands of items, we're doing this in two sweeps. A quick pass for the low-hanging fruit, the really obvious stuff in that kill.

But as we go through the kill, we'll also be marking things with an amber. So those are things that we'll review and come back in a second sweep. And one mantra we use internally is we're going to automate the facts, not the decision. So we are going to use, we hope, AI tools. We're already developing them that will help identify things.

So for example, our partner at 365, they're developing a tool for us that will not just look at the content, but we'll look at things like the use of logos. We've had a brand change three years ago, three, four years ago. So things that use the old brand will be flagged for kill. Equally, when people write copyright 2020 on the document, know that it might have been updated, but that's another flag for us that it's not an active document if it's got old copyright or regulatory information on it. So that AI tool will also pick up and flag that.

But as I say, when automate those kind of facts, they flag it to us, not the decisions. We're absolutely still going to have humans involved. Unfortunately for them, kind of making the final decision. I skipped over this slightly, but just to show we're already doing those quick wins. So we did the investment section.

You saw one of the big sites. We looked at one of the page libraries in that kind of set of sites. Basically found of the 316 documents, 203 just using the quick wins could be archived

straight away. That's 64% straight away that were candidates for deletion. And that one was without all of the AI tools.

That was literally someone just going down and checking. And we have 113 documents that we're now going to go through and do a second pass on. So really easy with a clearly laid out kill keep combined kind of framework set of documents with kind of what we class to get rid of documents without having to get into too much detail. And you know, it's quite funny. People often avoid taking ownership until you threaten to delete their page.

And then suddenly everyone is deeply attached to their content, even if it hasn't been touched ever. So be ready for those conversations. And I'll probably mention that again shortly. Looking at some of our other work streams, I mentioned governance. Absolutely really, really important for me is the governance, even though it might not be my forte.

Cleaning up the content is half the battle. Keeping it clean is the other half. And that's where it'll stop the sprawl and prevent us from doing it again in a few years time. So some key principles for our governance. Owners for everything.

We want zero orphaned content. Every page, every site, every document will have someone's name attached to it. Why? Because if you can't name who's responsible for a piece of content, you can't confidently say it's accurate or needed. And like I said, if nobody owns it, nobody fixes it.

And that flows in slightly to the second point on this. Clear roles and accountability. We're making sure everyone knows their job. And there's a bit of semantics in this. And we've had some challenges with semantics in the business.

We want to make a distinguish, a line between what is a content editor, the person who updates pages, and a content owner, the person who decides what's published, approves it, whether it can be retired. And we have a challenge where SMEs will often keep content up to date, but they don't want to own it. And so we're being really clear about what those different roles is, as well as the role of my central team to provide that support and oversight, not to control and do everything for everyone. So we want to balance that local ownership with central standards. We're building content lifecycle.

We've already got, like I say, some approval workflows. So we're just going to build on those a little bit more. So everything requires at least an annual review. If nobody reviews it, it flags up. We want to make that more routine.

So right now, it flags them three or four times. Then it comes to my team and we manually kind of harass them. We decided we're just going to archive that content if it doesn't get reviewed using kind of Power Automate. We're going to reintroduce our standards and training. We used to do a half day kind of intranet editor introduction course.

We found that we do that. And then two weeks later, they'll be calling us up with the same questions because it was quite intense. So we're splitting everything down, very similar to

before, into modules that are available on our L&D platform. And people will need to re-accredit every year. They won't need to necessarily do the whole course every year, but they will need to re-accredit.

And we'll also then pick up some of those changes that Microsoft might make or we might make to our design system. And as I said, our intranet editor community is about 500 plus people. So just to note as well, we use Viva Engage quite heavily here. So we have a community on Viva Engage just for our intranet editors. And we spend every week, we do Mick, one of my team, Mick's kind of Monday masterclass, essentially office hours, community call.

He's just there online for editors to join and chat to for a couple of hours. We post latest updates. Right now, my team are running the monthly kind of user research call of all the user research they've done and the analytics reviews. We then put that straight into that community and highlight some big things for them that's key. And actually that's been a real great way to get engagement because when they see we're auditing their content, they quickly want to jump in and respond.

And we're backing that community up with some steer co's, a governance forum with reps from around the business with executive sponsorship. Someone at the top of the business, someone senior who can help and block things and kind of own it aside from me as maybe the technical owner. And those technical improvements I mentioned, key tech pieces we're going to be doing. So search tuning. So we're under the hood of SharePoint search.

We're adjusting the kind of configuration, but we're also doing things like using the metadata against every page and synonyms essentially against every page that we've put in the metadata to create the best bets on SharePoint. So really simple tool you can do when someone searches for 1234, show document 1234. And ironically in SJP that's exactly what we have. A lot of people have got used to calling documents SJP 1234 rather than the name of the document. So they type that in quite a lot.

Those kind of random numbers appear high in our search results. So we're making sure every single one of those numbers links back to its document. So even though the document number doesn't exist anymore in the document, it's not in the title, we're adding it in the metadata and ensuring it's presented for some of those more legacy partners who still use those numbers. We're going to our information architecture is something we've done a few times. We've never drastically verged away from it, but we want to make sure it's iterative that we constantly look back using card sorting and other elements to make sure we understand our information architecture and we hold to it.

A lot of people say oh can we have a new site please? Sometimes with these rules around why that's right or wrong, but if we decide it's right we want clear rules in the information architecture about where that site might sit and which kind of area of the business it belongs to. And we want to make sure that's not hierarchical area of the business. We often have a challenge where three or four teams all own products. We don't want three or four teams

arguing over having separate sites for their products.

We bring them together. That's how our information architecture will force. I've mentioned the automated workflow. So everything from creating the sites and making people follow the governance, archiving pages, approval processes, and I'm happy again you can connect with me on LinkedIn etc. Happy to share kind of the more technical side of what we're doing there.

But we're doing lots of stuff around automated workflows to ensure that our content doesn't sprawl. Actively using analytics monitoring. So we will again do content health, identifying stale pages, low usage, broken links early, so we can catch that sprawl and make sure before it gets out of hand we're on top of it. And finally AI before we kind of really get into the AI. We're using AI to assist with oversight.

So it's great thing you know but this might only be a thing that works as kind of the UK audience. I don't know how your roads are for the rest kind of in the EMEA region but in the UK they're certainly not great right now. And so what we've said is Copilot isn't the pothole repair crew. It isn't the thing that's going to go out there and solve our content issue. It's the person that goes out in advance and paints those strange white boxes around the pothole that we can then sit there for six weeks to say here's a pothole that needs fixing.

That's how we're using AI. It will go through as I said it will help. It's helping us through various agents spot those challenges, spray that white box, flag it back to us so that a human can go then and make the proactive decision. Obviously if something is really clear red kill it's a quick decision to make but we just want to make sure that we're using tech as a force multiplier but noting that that kind of technology alone isn't the silver bullet and that's why things like the governance and the content audit are important. And the title of this talk before AI makes it worse so let's address that.

The reality is AI will amplify whatever state your content is. If we turned on Copilot or you know also use Claude today on our intranet it would be drawing from cluttered knowledge base. It would present answers confidently. It would be a great product but that's a terrifying prospect. So garbage in, garbage out.

So we've seen impressive demos of AI tools you know we know they're exciting we know it's critical to remember that AI doesn't automatically know what's right. It learns from what you give it and that's why we're talking about as well AI readiness. Whether we are AI ready for content, whether we're governance ready for that content, whether we have clean well structured authoritative content to feed these tools. Yeah think about it as you know looking after your kid. If you feed them a great diet then they'll grow healthy.

Feed them junk and maybe not so much. On the flip side we're optimistic about using AI. I'm probably the biggest champion of it certainly in my area of the business. Absolutely waving the flag for how it can be a tool of force multiplier for us. So yeah absolutely we're going to use it.

We're going to make sure that we use it to help with our cleanup and then we're going to use it going forward to make sure that we don't hit this again. So key takeaways for me build yourself a structure. Keep, Kill, Combine for us. It gives you a way to start. It's the hardest part with your content audit but having a structure like keepkill combined turns a very vague problem we've got too much content into something very practical.

You can look at a page and just say do we keep it? Do we get rid of it? Does it mean to merge? Simple decisions gets you moving. Governance matters probably more so than the cleanup.

The cleanup feels like work. You know it's very hard. It's very visible but without the ownership review cycles extra you're going to be back here in two or three years having the same conversation again. So the shift is not just cleaning the content but deciding who owns it, what happens to it because like I said if you don't name the owner you can't trust the content. Less clutter is not a compromise.

A lot of people go well I need to make sure everything is on the intranet. I think there's a mindset you know more content equals more value. In the practice I think we all know it's opposite. Every duplicate outdated piece of content is an extra bit of friction for your end users. So when you remove half the content from an area you haven't lost things you've made it far more usable for your audience.

AI readiness I've already mentioned just in the last slide but a lot of excitement about Copilot definitely for me but they don't fix messy content it scales the messy content. So question yourself are we ready for AI? Is our content clean enough that gives us the right answers? If we have a maternity policy or a you know family parental leave policy in one geography and a different one in another one what happens with that content? How will it give the right answer in the end?

Can we trust the answers it gives? Clean content is what will make AI useful not the kind of other way around. And finally for us it's you know it's always an ongoing piece of work. There isn't a point where you can finish this and move on. Content will change, the business will change, people move roles.

So for us you know using things like our champion network, our editor community, the weekly office hours, the having great business relationships, regular analytics, constantly talking about content quality is important. And on building relationships one kind of final point from me people won't always want to own content as I said earlier but they will fight very hard to keep it especially if their little face is on it somewhere. So a lot of this is as much about the people as it is about the process. Having those relationships with those stakeholders around the business is absolutely core and central to making a project like this work. So yeah that's me talk fast to get through it all but if you want to see how this plays out the next year feel free to get with me on LinkedIn or elsewhere and absolutely free to answer any questions.