

The Open Source Way

Episode 16: The Open Documentation Initiative – The Gold Standard of Collaborative Documentation



Transcript

Karsten: Welcome to the Open Source Way, our podcast series, SAP's podcast series about the difference that open source can be. In each episode we will talk with experts about open source and why they do it the open-source way. I'm your host Karsten Hohage, and in this episode, I'm going to talk to DJ Adams and Jens Haley about the Open Documentation Initiative. Hi.

DJ: Hello.

Jens: Hi.

Karsten: Hi DJ, hi Jens, nice to have you here. Jens has been with SAP for more than 20 years, and after working as an author at SAP, he coordinated the delivery of bigger documentation projects. He is the workstream lead for the technical implementation of the Open Documentation Initiative. DJ started working with SAP software in the 80s when everything was still written in assembly language, the GUI was a green screen, and the overnight batch job schedule was massively important. He otherwise studied Ancient Greek, Latin, and philology at university. He's written books for O'Reilly and SAP press and has done all sorts of jobs in his life; as a body collector, a corpse dresser for a funeral company, a processor of animal waste products – wow, you've got some morbidity going there – a nightclub barman, a chauffeur, you name it. He's a Developer Advocate at SAP and Community Lead for the Open Documentation Initiative. DJ, what is a corpse dresser?

DJ: When at funerals, the deceased ought to be viewed in ceremonies and churches, whatever, with the lid open, you want to make sure they look great with the clothes that the families provided. And so that's what I used to do.

Karsten: Okay. Is it not a kind of a scary job to do?

DJ: Not scary, no, it's sad. It's very poignant. But overall, no, I mean, it's just part of, you know, part of life's rich tapestry. It's part of life and you get to meet all sorts of amazing people in the trade. It was a very memorable time.

Karsten: Yeah, I guess "scary" just comes from, I don't know, watching the wrong movies or something; I always expect them to start talking to me while I'm dressing them or something.

DJ: No.

Karsten: No, okay, never happened. Jens, I know about you – and, I had to think about a movie beginning when talking about corpse dressing, I don't know if either of you know the Big Chill. It starts with that, it starts with...

DJ: Oh, yeah.

Karsten: The opening scene is basically a corpse being dressed, which at first you don't know. But let's not talk about dead people much longer. Besides what you do for SAP, you enjoy photography. Is that right?

Jens: Yeah, I mean, it depends how much time I have, but I was always interested in photography and whenever I have the time, I enjoy doing it.

Karsten: And in a special direction? Do you do landscapes, portraits, I don't know, special cameras or whatever?

Jens: Well, at the moment, I got into macro photography, so I don't have to walk fast, I can just...

Karsten: The real extreme macro, where you can only see the eye of the snail or something?

Jens: No, no, not that close. But I mean, it's really a challenge to have the motive sharp. So, I work with this focus stacking technique. That's something I tried out. Yeah. So, I'm a beginner there.

Karsten: Yup. I always find that I probably need to get into that focus stacking to make any sense out of macro. I do a bit of photography myself and macro, I totally confess I haven't conquered yet, really. But that's not what we're here to talk about. DJ, can you enlighten us? What is the Open Documentation Initiative?

DJ: I'll try and enlighten, folks. In a nutshell, it's an initiative to make SAP's documentation collaborative. I think one of the things that I, over the decades, have been really, you know, happy to observe, is, you know, SAP's gradually moving towards an increasing openness, right? You know, milestones like the introduction of open protocols like HTTP in the early days. Many, many ongoing contributions to open-source projects out there. The adoption of OData as an open data protocol. We've got our Open Source Program Office that's been established, Open Source in UI5, you know, there are so many different things, so many different steps that I've seen that SAP has taken. And this is, to me, you know, the next step. It's basically, you know, feedback on documentation has been, for the most part, a one-way street. You know, there have been facilities to provide feedback, but for me, it's a little bit sort of like talking into a box, but there's not really much response. And the Open Documentation Initiative makes the whole collaborative process interactive. So, for me that's the key thing; collaboration and interactivity. We've got one pilot live at the moment. You know, this is still really early stages. We've got one pilot live, which is the documentation for the business application studio, and we've also got a meta pilot live, the contribution guidelines themselves that explain how to contribute. They're also subject to this Open Documentation Initiative. We're looking at more pilots going live soon, so obviously, watch this space. It's important to say that, you know, it's not all of help.sap.com that's being open for contributions. We're focusing primarily, with these pilots, first of all, on SAP's Business Technology Platform content.

Karsten: Okay, just to make sure I get that right. We get feedback here in this podcast, often from open-source projects that, of course, one of the important parts to collaborate with the community is documenting the open-source project. What exactly is the difference between the Open Documentation Initiative and these open-source projects?

DJ: That's a super question, actually. Maybe it's worth turning it around, I mean, first of all, one difference really is that the primary goal of an open-source project is to

collaboratively produce software, whereas the primary goal or the vehicle of what the Open Documentation Initiative is all about is the official standard gold standard, let's call it, documentation, that has traditionally been produced sort of by internal SAP resources. You know, we're now opening that up for collaboration by external folks because, you know, I joined SAP only four years ago. So, 30 years before that, I worked in SAP space the whole time, and I know that so many people out there have got so much experience and so much to contribute to documentation. You know, they know what works, they know what doesn't work. And so, this, for me, is, you know, it's an obvious thing to do, right? So, in a way, one of the differences is that, you know, we're talking about basically standard documentation rather than sort of open source or rather than documentation that is connected to a part of open-source projects, not repos on GitHub, for example.

Karsten: Okay, so it's the regular help.sap.com documentation we're talking about, not something of some back corner GitHub or very prominent GitHub project, right? Okay.

DJ: Nice save there, Karsten.

Karsten: Yeah. Let's quickly forget that. No, but it's also not like with some online shops or whatever that we're outsourcing the writing of product information to our customers, is it?

DJ: No, it's definitely not, although, I mean, you know, I think one core similarity is that it's about conversation. One thing that we may get onto is that the contribution process, whether that's through feedback or contributing actual content suggestions, yes, those are the vehicles. But it's all about the conversation. It's all about the conversation between internal folks and external folks. Just like, you know, the conversation on, you know, a large retailer online is about the conversation between the reviewers. However, you know, that's not always a bad thing anyway, right? Look at the SAP community. It's amazing, different folks helping other different folks, you know, so that is a good thing.

Karsten: Okay. Speaking of contributions and discussion and everything, how has the activity in ways of contributions been so far?

DJ: I would say one of my favorite phrases: "It's small but perfectly formed", you know, we're starting out slowly and we've got, like I say, we've got the business application studio documentation and we've had some great suggestions through feedback. So just to make it clear, there are two ways to collaborate, to contribute. You can provide feedback, which is "I think this" or "I've observed that" or "How about this?", and there's content contribution where you say: "Oh, well, you know, I saw this thing and there was something missing, or something needed changing. So here is that missing thing here. Is that changed?" Folks who are listening who were aware of, you know, the GitHub based collaboration process, will recognize those two things as issues and pull requests, right? And so, we've had a mixture of both. Fixing small things like typos, bad URLs or whatever, but also making suggestions and contributing new bits of information, maybe also reordering things to make things clearer. So, all sorts of different wonderful contributions.

Karsten: Okay, I guess we have the general idea now. What further goals, what things on the roadmap does the project have, in a little more detail, soon?

Jens: Yeah, maybe just to add to DJ, I mean, first of all, I think the main goal is to bring everybody together in this external collaboration process. So, it's not only about authors being able to communicate with the outside world, but also to include our SAP developers. I mean, they are part of the documentation creation process, but it's really great if they can see through the feedback that we get on the documentation, you know, where the pain points are in the software, also. But I would also like to elaborate a little bit on, you know, what's not really the goal. So, I mean, up to now, we have just talked about the external view, so that people may be on the help portal. They see something they would like to correct, or they see an issue and they can directly navigate, by the way, from the help portal to GitHub.com, so, we have updated the UI of the help portal and there's an edit button that then immediately opens the pull request on GitHub.com, so, that's the external view, but that doesn't mean, really, that the source of the documentation is on or must be on GitHub.com. So, we use markdown to have this collaboration process going, but authors at SAP do not actually use markdown, or not all of them, and the main system to produce content is the DITA CMS. And then sometimes, there's this reflex, you know, when we have this project and they see it, or we have external collaboration for GitHub to say, well, let's use GitHub as the source for the documentation. But I would say, at the moment, we are using a side-by-side

approach, so we still use our content management system that we have, our DITA system, and then we use GitHub just for the collaboration aspect. And of course, we have teams that also use GitHub to develop code at SAP before, you know, it's released somewhere. But with regards to documentation, I would say we are still in an exploration phase. So first of all, that documentation has to go to the help portal as well. And yeah, we're looking into the details, you know, because GitHub is not really a content management system to me, it's a versioning system. And what I have seen is that whenever people start to use GitHub as the source, they start to build a content management system. And so, this project is also a way to see, you know, how we can get to a standardized approach there.

Karsten: So, I think you share my perception that I have developed in the fate that basically I have at SAP; I work as a non-developer among all developers and software architects in the organization that I'm in. And they always try to make me work on natural language documents using GitHub. And I always say, well, no, it's not really for natural language documents, it's for the branching and versioning and everything of code, but it's not really 100 percent the same mechanisms. Is that what you also experience, Jens?

Jens: Well, I mean, I would say, you know, people choose the tool that they like and, of course, developers like GitHub, and that's what they work with every day. And then, you know, just as an example: I was part of the Gardener team as an author, and Gardener is also open source. So, they used GitHub a lot. And first I was a bit skeptical, you know, because I hadn't used GitHub that much before and I was hesitant. You know, why do I need to use it? You know, we have our content management system and so on. But then I also really saw the benefits because I could really do good reviews with people that were not located in Germany. Instead, you know, they were in Sofia or wherever, and we could really go through my documentation line by line if we wanted to and review it, you know, as if we sat in a coffee corner. So, I definitely see that, you know, that's the big plus point of GitHub, those collaboration capabilities. But when it comes to documentation, there's more, you know, people want to have a proper navigation for the documentation. We don't want to be lost in correcting links all the time. If you make structural changes and whenever people enter that domain, they soon realize GitHub might not be the right tool or, you know, they use some other open-source component to pimp it up. Yeah, and I'm really excited about where this whole

story is going. I have a feeling it will still be a side-by-side approach and we will see, you know, when GitHub is the source, where the limits are and what we can support in future.

Karsten: Just out of personal interest, then when you said you were so very well able to do cross-location reviews via GitHub, did you just use it online, like in the web interface, or with all these offline syncing tools?

Jens: No, I also used those offline syncing tools because it's a lot easier, you know, if you clone the content and then you have the editor that you like. I mean, I personally use vs code. And there are a few things that already support authors to locally work on content, for example. I don't know whether you've heard of Acrolinx, which is a tool to kind of check the language, and we can actually use that at SAP together with markdown, so we can connect from this code to our Acrolinx server. And instead of checking our DITA content, which is an XML standard, we can also apply the same checks to markdown. So, I mean, that works well.

Karsten: I know, I was, at some point, even asked for my opinion and some testing of Acrolinx, and I personally feel that it contains the danger of people relying on that for their language check rather than having a four or six or eight-eye principle. And it can never do what a four or six or eight-eye principle can do. But anyway, this is not what we're here to discuss in detail, the AI versus the human control and quality management of text. How do you expect the Open Documentation Initiative to continue? Where should it, in your opinion, lead?

DJ: Well, I mean, actually, I'm struggling with not saying anything about Acrolinx, because I have my own experiences as well, but anyway, let's get on to answering that question. As I mentioned before, we're focusing right now on Business Technology Platform topics, and we've got some pilots, more sort of developer tech focused pilot documentation, that's lined up. So those will be sort of appearing at some stage. You know, I can't really say when, but I think also I think it's fair to say that GitHub itself will be in some way or another, you know, that will become an integral part of SAP's documentation infrastructure. I think it's important to, I mean, it was such a great set of points that you were talking about before. And I think it's worth at this juncture, just being clear to the listeners, you know, we have got three different things going on here.

We've got the distributed version control system, which is Git, not GitHub. GitHub, for me as a developer, but also as a natural language author, that's not what version control is all about. That's about collaboration, that's about working together. You know, it's got amazing workflows and everything, that's what Jens was referring to. And then we got a content management system. And GitHub is not a CMS, you know, DITA is our CMS. So, I think it's, you know, these things are very closely connected, but it's also important to be able to sort of separate them in one's brain, right?

Karsten: I just had to smile because what our listeners don't know is that before the recording, we talked about the early days of the internet, and DJ pointed out that, of course, the Internet is not the same thing as the Web. And now, you're basically doing the same thing and saying GitHub is, of course, not the same thing as Git.

DJ: No, it's completely different.

Karsten: That's what made me smile, that you're like, very keen on pointing out these semantic differences, basically.

DJ: Yes.

Karsten: And anyway, I just remembered something else as we were talking. That was, at some point, as we were also speaking of the past of the Internet and everything wasn't there; at some point, something called SAP Docupedia that, from the basic idea, was something similar, like a collaborative writing of SAP information, at least. Is that still around and how is it different? Or is it gone and how is it different? Jens, maybe?

Jens: I must be a little bit careful here because that was one of the projects of my former manager.

Karsten: Aha!

Jens: But what I can say is, no, it's not around. I think one of the problems of Docupedia was that it was tricky to find out when, you know, everybody could collaboratively work on documentation. What is the actual official SAP version of it? That was the main challenge there, and I think they couldn't figure that out. And now

with GitHub, there's no doubt, you know, that whatever people see on the help portal is still the official SAP documentation, and we have found a way of kind of gatekeeping, you know, what gets into it. We use those pull request reviews that can only be approved and merged by SAP employees. So yeah, maybe we learned from Docupedia, but, you know, it's not a follow-up project or anything. But yes, it was a similar idea.

Karsten: Okay, so it's not the Wiki-type, totally open collaboration platform, but, basically, the veto, the approval roll is all still funneled, tunneled through SAP approving of changes and feedback, right?

Jens: Yes.

Karsten: Okay, great. But to jump back to where it should lead: Is the idea, even if it's not like Docupedia, is the idea in the end that on every page on the SAP help portal you basically have a button that says, "Suggest change"?

Jens: Well, I don't think so. I mean, the help portal not only contains documentation for technical people, but we have end users that use our marketing tools, we have administrators, but I'm not really sure yet, you know, whether they would have a GitHub account, so for the absolute end users. I would say it's probably not an option to say, well, you have to have a GitHub account to be able to provide feedback. That would be a bit over the top. But I mean, for developers, it's only natural, you know, to use the tool that they already use daily anyways to provide feedback to us and to get into a conversation.

Karsten: So, you would basically, by the choice of tools and the way of user interaction, implicitly limit those who will provide feedback down to those who know what they're doing.

Jens: Well, I mean, maybe it sounds a bit ridiculous, but you have to see that at least at the moment, you know, every documentation that we bring live to GitHub is also like a mini project. And we also need the resources to do the whole thing. So, it's not as if we could push a button and then all the documentation of the help portal is suddenly on GitHub.com. But so, I mean, that's why we choose, you know, where it doesn't make

sense to have this external collaboration process for developers, mainly. So, we have a look at developer documentation.

Karsten: Okay. Okay, I'm just kind of trying to make some things clear here so that, I mean, this is going to be a publicly available podcast, so we have to make sure that not everyone expects that they can send us these millions of feedback messages on any page of documentation that they always disliked or something, but that it's still a rather targeted approach to certain areas.

Jens: Yeah, I mean, maybe just to add to that, I mean, if people would like to give an overview, you know, where that is possible, because it will not be that easy to find, you know, the documentation, where you can do it by looking at the help portal. But all the documentation that we include in this collaboration process is found in our SAP-Docs organization. So, when you go and GitHub.com...

Karsten: Shortly, shortly, before we finish, we'll come to how people get involved, I think. I think you have something to add at this point here.

DJ: Yeah, I was jumping up and down, sort of in audio mode, but that was such a great question, Karsten, you know, are you saying that we're restricting the feedback to people who know what they're doing. I love that, such a brilliant question. Harsh but fair. That's one of my favorite phrases as well, and I think I'm old enough and daft enough to say, well, yeah, obviously, because you know, first of all, yeah, you need to know what you're doing. It's not difficult. It's not rocket science, but also going back to the discussion of, you know, Docupedia versus the official help.sap.com documentation. You know, we're dealing with the crown jewels here. You know, it's not just free for all – Wikis are amazing, right? I'm a big Wiki fan – but this is not a Wiki, you know, SAP are opening up the documentation, the official gold standard documentation. So, there has to be, in my opinion, some level of review and standard that's easy to achieve and we described in the contribution guidelines. Otherwise, you know, it's going to end up in chaos, right? So, I think it's a nice balance between, you know, sort yourself out and contribute to this wonderful thing.

Karsten: When I asked that question, by the way, I had to also think of parts of your CV, because if you studied ancient languages and philology and philosophy and so on,

and some of these texts also sometimes are written with an attitude of I only want people to understand this who really understand it. At least that's what some philosophers sometimes claim that they intentionally don't write it easy. But okay, no, we're not blocking anyone out, this is not what we're saying. We're just saying this is a dedicated approach and not like a public throw at everyone. Right?

DJ: Exactly. There's an entry cost to have a conversation, and that entry cost is to be able to have a conversation where you know who you're talking to and have that thread and that entry cost is the collaboration, you know, an ID on the collaboration platform that we're using. That's it.

Karsten: Yes, that was very well put. I will not say any more before I screw it up again. Anyway, I just mentioned before, as we're kind of approaching the end of our time here, what will people that now are interested, like within SAP and outside SAP? Where do they get started with learning more or with even contributing?

DJ: I'm going to give a couple of pointers, there's a really fun video starring the one and only Jens, my wonderful colleague on this project, and I'm in there as well. So, it's about, I don't know, it's about five minutes long. It gives you an overview of what it's all about. I think it's great, you know, regardless of whether I'm in it or not. So, we'll put a link to that, I hope. And there's also a blog post that we wrote and published when we first kicked off the initiative earlier this year, and that gives you all sorts of details about how to collaborate, what the different types of collaboration are and so on.

Karsten: Okay, cool, I'm sure we're going to have all the links to the materials just mentioned under the podcast information once it's published. Now, apart from that, are there any other key takeaways you'd like all our listeners to remember, like maybe you take turns and maybe Jens, you start.

Jens: My key takeaway would be that GitHub does not replace the CMS that SAP uses, but it complements it for collaboration.

Karsten: Okay, DJ?

DJ: I would say this is just another step on SAP's journey towards openness. And we should all embrace it.

Karsten: Great. Then we'll just call it a day and I'll say goodbye. Thank you very much again, Jens and DJ.

DJ: Thanks for having me.

Karsten: Thank you all for listening to the Open Source Way. If you enjoyed this episode, please share it and don't miss our next episode, published on every last Wednesday of the month. You'll find us on openSAP and in most places where you find all your other podcasts, like Apple Podcasts, Spotify, and the like. Thanks again for listening and bye. And Miguel, you can stop the recording now.