The Open Source Way

Episode 02 - Corona-Warn-App: Behind the scenes

SAP Open Source

Karsten Hohage: Welcome to The Open Source Way. This is SAP's podcast series in which we'll talk about the difference that open source can make. In each episode we'll talk to a different expert and we'll talk to them about why they do it the open source way. I'm your host Karsten Hohage and in this episode, I'm going to talk to Sebastian Wolf, or for our international listeners, Sebastian Wolf, from the SAP Open Source Program Office. And he was especially involved in building the Corona-Warn-App for Germany. Should I call you Sebastian or Sebastian during our talk?

Sebastian Wolf: Sebastian is ok. Whatever you please.

Karsten Hohage: Was that the German version or the English version? I couldn't really make that out.

Sebastian: But also, you were something in between. Anyway.

Karsten Hohage: Yeah, I'll probably do something in between anyway. Not being an English native speaker myself.

Sebastian Wolf: Same for me. I'm also not an English native speaker.

Karsten Hohage: Sebastian has been a development architect with SAP for quite a while now. He actually started with SAP in 2003. Back then, not as an architect but as a student still. And he has been in different development roles for ABAP development tools, for supplier relationship management and since 2020, he's part of the Open Source Program Office and in this role he has taken care of the Corona-Warn-App project in a certain role that we will learn from Sebastian soon. But before we learn about that, Sebastian, in this entire Corona situation, are you in the office or I seem to remember seeing you in our last video call that you were sitting next to the laundry dryer. That was not really the office, was it?

Sebastian Wolf: No, no, that wasn't the office, and yeah, thanks for welcoming me here as well. I'm sitting at home either in the room with the laundry dryer or on my sofa as it is today. So I think it's now half a year more or less since March 13th. It was actually Friday 13th.

Karsten Hohage: Actually, I do remember that as well, Friday 13th was the day when SAP finally said you're all staying home now, right?

Sebastian Wolf: Precisely.

Karsten Hohage: And I think five days before it was everybody who doesn't have a reason stays home and five days before that, everybody who has a reason stays home or something like that. Yeah, the bad thing is that you are not sitting with the laundry dryer that we cannot talk or yes, we can, because our listeners do not see us anyway. But with a laundry dryer, you also have quite an army of stuffed animals sitting with you, right.

Sebastian Wolf: Yes, I do. So, yeah, on top of my head, behind me of course, you would under normal circumstances see Tux the Linux penguin. Yeah, which is with me for more or less 20 years now. And now, finally working here in the Open Source Program Office. But of course, I've been an open source enthusiast, Linux enthusiast for quite a while. So I've made my hobby, so to say, a normal day job.

Karsten Hohage: That sounds like what everyone basically dreams of. I would have to think about what I'd be if my hobby was my job. My hobby has at least not always been making podcasts, but I like doing that. Of course, I like doing that. But let's not get lost into kind of collateral topics here. As we're talking about the Corona-Warn-App for Germany, which is only rolled out in Germany, for all I know at least, can you just briefly explain what it does?

Sebastian Wolf: Yes, sure. The Corona-Warn-App is like a diary that, in a more or less anonymous way, writes a log of all your contacts. You can then, in case you are infected, just notify those contacts you have had in the last 14 days based on this log. There's also a huge infrastructure behind to, let's say, also check these anonymous exposure data to have a connection to the German health infrastructure because all this testing, the laboratories also need to be connected to this whole infrastructure. That's it more or less...so on the one hand there is exposure logging and tracing and on the other hand, the connectivity to the German health landscape.

Karsten Hohage: It notifies me if I had contact with someone who is later on diagnosed with the corona virus. So what was in that project, that you were a part of, what was your role in that?

Sebastian Wolf: I had several roles, or mainly two, so to say. The most prominent one was to kick-start the open source workstream. One of the central aspects was, also due to public debate, which we will also come to a little later, probably…

Karsten Hohage: I guess so!

Sebastian Wolf: …the whole infrastructure needs to be developed out in the open as open source so that everybody who is interested can really check - okay, we are not doing anything against the community, against the German people. And of course, also to enable other people to contribute, both with issues but also bug fixes and probably additional features. So that was the one aspect. And the second aspect was community management. Of course, once we open-sourced documentation, coding and related things, people also engaged with us. And we needed to check that everything is doing in a good way so that we also need to enforce a code of conduct, for example. So that we have a healthy discussion, a productive discussion, and that everything is done in a professional way. And we don't, let's say, let things go into a direction that might not be beneficial for the case.

Karsten Hohage: OK. So that is basically the usual job of a community manager, the last part that you described, only that it had this special parameter of being about a pandemic that concerns everyone and everyone has their own ideas about how things are to be handled and everything, which was probably a little more challenging for the community manager, right?

Sebastian Wolf: In some discussions, definitely so, of course, the whole discussions and also concerns were reflected in some of the issues we got on GitHub. But I was also positively surprised, to be honest, that we didn't get that many people who questioned the role of the Corona-Warn-App in general. So in general, we had a pretty, pretty solid discussion, very professional.

Karsten Hohage: One thing I wondered about when you were just answering the last question is, was the main reason to do it open source really the transparency or was it getting the contributions?

Sebastian Wolf: It was definitely the transparency.

Karsten Hohage: OK.

Sebastian Wolf: So you probably also remember the public debate here in Germany. There were several initiatives already starting before we even got the contract from the German Federal Government. Some of them were not really developed in the open. And we also have, of course, a very, very good community especially when it comes to data privacy, also related to the German history and what has happened over the last decades. They were really, really urging the German Federal Government to make this open-sourced because it needs trust. And without trust, people won't use this Corona-Warn-App.Transparency was the main driver to make that open source.

Karsten Hohage: I don't even know what you're saying about the German history, what are you trying to implicate there were times when Germans spied on Germans.

Sebastian Wolf: Yes, I think so, yes, that's something I want to…

Karsten Hohage: Yeah, I'm just kidding. If you're listening from a very far away part of the world and you're not very old yet, you can look it up. We were the best professionals on spying on each other and that is why transparency was very important here. You were the community manager and the public relations guy in this, right?

Sebastian Wolf: One of them.

Karsten Hohage: How did you end up in that role?

Sebastian Wolf: At the very beginning, we were simply contacted by the people within SAP, who started this Corona-Warn-App project, because we have an established open source process within SAP. So each and every project that wants to go open source or that needs to go open source needs to complete the so-called open source outbound

process by SAP. There are some legal obligations: we need to check the license, we need to check if the responsibilities are fine. And so the Corona-Warn-App also came to our desk. It was not the usual project, so it wasn't done by checking: "Ok, we use this and that license. Of course, responsibilities are clear, deadlines are clear, repositories and so on and so on". They needed much, much more. Also with respect to know-how, knowledge and some background – who already has some experience with open source communities, either from SAP itself or in private life.

Karsten Hohage: Where did you gain yours?

Sebastian Wolf: I'll come to that in a minute.

Karsten Hohage: OK, sorry.

Sebastian Wolf: I basically volunteered for that and the role simply grew over weeks. And in the end, I was one of several people. Of course, I wasn't the only one who basically did the community management and answered the questions and also checked the website. Coming back to your question, where did I gain mine? Of course, I haven't worked in such a huge open source project so far. In my private life, developing software on mobile applications – not for iOS and also not for Android – it's for a niche mobile operating system called Sailfish OS. They have developed several apps, also dealt, of course, with the community there. And it's on a much, much lower scale than here. But at least I had some experience how to deal also with very, very critical people.

Karsten Hohage: So that at least means you did have a public GitHub profile before you became the community manager for the App.

Sebastian Wolf: Indeed, I even had two, to be honest, so one business and one for my private projects.

Karsten Hohage: OK.

Karsten Hohage: I think you have a pretty particular nickname you told me before. Do you want to share that what your private GitHub nickname is?

Sebastian Wolf: The private nickname is Wunderfitz. So even for many, many Germans, it might not be obvious at the very beginning what that means. It's coming, it's a word from my local dialect where I grew up – close to the Swiss border in very, very southern Germany. And it doesn't mean anything else but curious person.

Karsten Hohage: OK. Let's return to the Corona-Warn-App. In the end, basically, do you feel that the fact that we have a technologized world does help in dealing with such a virus?

Sebastian Wolf: I'm convinced that this is the case, absolutely. So if we haven't had mobile devices in our pockets equipped with Bluetooth, which enables us to at least estimate distance and close contacts with other people. Just because it's there, you don't need to buy additional hardware. You don't need to invent anything else. You only need to develop some additional apps on top of the existing devices. That, of course, facilitates that a lot.

Karsten Hohage: You have already started to share some details when you said Bluetooth because that's the basis on which the app works. Why is it Bluetooth? Another thing that comes to mind is GPS, why was the decision made for Bluetooth?

Sebastian Wolf: There was also some kind of an evaluation and learning process especially at the very beginning of the pandemic. There were several countries who were either exclusively or also partially using GPS, for example, as a means to trace back people if they have met or not. For example, Iceland, if I recall correctly. But due to the fact that GPS doesn't work properly indoors and it's also not that exact so that it really can trace you back if you have met with somebody, let's say, at a distance of one, two, three or four meters. Then it's not helping that much. It simply can only tell you, OK, you probably might have been in the area. Also, if you live in a four- or five-story house, it's also not really close. So it doesn't help that much. It could only be serving as an additional indicator, if you already remember: “Ok, that might be the case that I have met this particular person.” Bluetooth or Bluetooth low energy, what we're using right now, is much better because with respect to the intensity of the signals and also the way how it works, because it doesn't need a satellite. You only need the devices in your pockets. It also works indoors. And you can estimate – of course, it's not in the exact distance measurement – but that's also very important. It's just, let's say, based on

several measurements that have taken place also with the Fraunhofer Institute here in Germany. So this and that attenuation. The signal is not that strengthy anymore. After one meter, after two meters, you can derive: „OK, this particular person is now in a distance of approximately, very important, one meter, two meters or three meters away from me.“

Karsten Hohage: Well, I guess the virus is also, or at least the infection with the virus, not exact.

Sebastian Wolf: Precisely.

Karsten Hohage: It doesn't exactly infect you within 1.5 meters and doesn't infect you anymore after 3.2 meters or something like that.

Sebastian Wolf: Something like that.

Karsten Hohage: Totally fine. Isn't also another problem with GPS, wouldn't that have required some centralized architecture, if one had gone with GPS?

Sebastian Wolf: Not necessarily. So in the end, also GPS could have been done in a decentral way so that your only records are the respective locations on your local device. And only afterwards, if you have been tested positive, just upload your, let's say, location history to the server. However, it's much more privacy-problematic because you would basically expose your location history. With Bluetooth, you only expose anonymous or pseudonymous IDs which are random. So you couldn’t even trace back the locations of these let's say exposures, where you met people. So it's much, much better from a privacy perspective. You could also do the Bluetooth technology in a central and decentralized way. So there are several dimensions to this problem. That's also everything what I wanted to say. You could use GPS in a central and decentralized way from a privacy perspective. The decentralized Bluetooth way, it's the best from what we know and what also works, by the way.

Karsten Hohage: Correct me if I'm wrong. All the apps – before you maybe get notified by your doctor or by the clinic or someone that you have met is tested positive,

everything that happens until then is only that the device is basically peer-to-peer exchange: "We have been close to each other".

Sebastian Wolf: Exactly. They exchange random IDs, and everything is stored exclusively locally on your personal device. Nothing is shared with the central infrastructure. And if you have tested positive, you have once again an opt-in that you share these local IDs that you've met over the last 14 days with a server. So there's even an additional decision point there if you want that or not.

Karsten Hohage: OK, and so basically the opt-in point one is, I install the app, opt-in point two is, I switch Bluetooth on and opt-in point three is, when I get a positive result, I actually report this back to the infrastructure.

Sebastian Wolf: To the infrastructure, exactly.

Karsten Hohage: OK.

Sebastian Wolf: And the infrastructure can then distribute the keys and notify the other people.

Karsten Hohage: And even when I report back to the infrastructure, what becomes known about me centrally or to other users?

Sebastian Wolf: The only thing that becomes known centrally are the random IDs of your device that have been sent out over the last 14 days.

Karsten Hohage: And it doesn't even know that this is me.

Sebastian Wolf: Exactly.

Karsten Hohage: OK, because I'm also only an anonymous key.

Sebastian Wolf: Several anonymous keys.

Sebastian Wolf: The fun fact is, it's even that good from a privacy perspective that we can only somehow infer how many people have actually been reporting back because the anonymous keys change from time to time. Of course, there are several attack points. There also have been several papers written about this one. Of course, it's always some kind of a compromise. I've been engaged with data privacy initiatives over the last year or you could already say decades. It's the best option from my perspective, that could have been done, in order to tackle that problem. Definitely.

Karsten Hohage: And I just wonder, are there timestamps on the proximity entries in your diary or are there are no timestamps?

Sebastian Wolf: They are timestamps. And also when two devices meet...

Karsten Hohage: Mm hmm.

Sebastian Wolf: …there's a time stamp, of course, and there's this attenuation value and some additional metadata but not that relevant. But these are basically the most important aspects in order to decide. OK, was this exposure relevant or not? Because you also need to say – was it for 5 minutes, was it for 10 minutes, was it for 30 minutes? And the longer, of course, the more problematic it might be. And the closer, so the lower the attenuation was, the closer the people or the devices have been together and the more likely it is that an infection has actually been taking place.

Karsten Hohage: Would timestamps not possibly enable me to re-engineer a location track as well?

Sebastian Wolf: Yeah, only if you basically put out Bluetooth receivers on several areas, for example, in a city. That was one of the possible attack points. People said: "OK, if I place several Bluetooth trackers in Berlin, in Munich, somewhere else and I could also probably couple it with a camera or something like that. I record all the beacons which were sent out and two weeks later I checked the public lock. Or have I seen some of them? Could I possibly trace them back to the camera recording or something? That was done, of course, completely illegally". That was one of the attack points. But of course, think about it, how likely it is and how much effort you need to put into that attack to make it work on a larger scale.

Karsten Hohage: If it involves putting up cameras in public places, like unauthorized cameras, we're not talking about actually a vulnerability of the app itself anymore.

Sebastian Wolf: Yeah, of course, you combine several data and if you can combine additional data to support that, but as I said.

Karsten Hohage: Yeah, OK. I mean, that's a major "big brother effort" there to achieve that, I guess. What else was critically seen in public?

Sebastian Wolf: The huge project in the end. Right. People said: "OK, why does it involve Deutsche Telekom and SAP? Why does it take so long and why is it not ready? It's such a simple thing". And to that I simply have to say, especially when it also became obvious, partially, at least from our discussion with data privacy, that a lot of things need to be taken into account. It's such a big project. Especially if you target the whole population of Germany or a certain percentage. It's becoming effective if it's used by 15 percent of the population. You need to check security infrastructure that it works 24/7. You need to have the legal aspect covered. You need to have the software development properly covered. Also, continuous maintenance and that's often forgotten, you also need to cover documentation and user-assistance in the end especially when it comes to end users. There are many, many questions. People simply want to know how it works and they simply want to call somebody. And we have a hotline in place, both for the pure usage of the app and to get the transaction number to upload your test results or your diagnosis keys. Diagnosis keys are the keys that are these random IDs which you have exchanged and you want to inform the other people and that all combined. Of course, I understand, or we all understand that it might not be obvious at the very beginning, why it definitely needs that many efforts, why it needs that many resources. But just put it the other way round. If that hadn't been done that way, probably there would also have been many, many discussions. OK, why was this aspect not covered? Why is the security hole still there? And so on. You all know these discussions. Also, given the timeframe we had that was simply the only chance to put some additional resources on top to make this work and to bring that to the public within just, let's say, 60 days or so.

Karsten Hohage: So in the end, those that brought up the criticism that they could have done it themselves with their start up or whatever, could have done it 10 times faster and 10 times cheaper just simply did not reckon with all the collateral activities basically, that coding, in the end, is maybe 10 percent of the effort or something like that.

Sebastian Wolf: Probably, I wouldn't phrase it to that extent but I simply want to say there's a lot more behind the curtains that need to be taken into account in such a huge project than just pure software development and probably some documentation and some community management.

Karsten Hohage: You just mentioned the topic of community management again which was actually your core responsibility in this one, right? Where there are a lot of contributions from people outside of SAP or Deutsche Telekom.

Sebastian Wolf: Yeah, so especially at the very beginning when we brought this out – also even before we actually released the code with respect to documentation and also related aspects such as a website with FAQ entries – we really had a lot of contributions. So that actually also surprised me. I talked to one of the developers or several developers and similar. Several of them had similar feedback. So for each and every hour, they needed to invest to review external contributions especially when it comes to quality aspects, the value of what they received from the community was at least double the effort they needed to put in. So for each and every single hour they needed to invest, they get two hours of value back. And that's really something we cannot underline more here. That's really the value of open source and contributions in such an open source project. However, and that's the other side as well, I also need to admit that, of course, we couldn't accept as many contributions as we probably wanted to get. And that's pretty much also due to the special nature of this project. So the special nature of this project is, we have on the one hand the institutions who contracted SAP and Deutsche Telekom to build that infrastructure, the app and everything around it. That's the Robert Koch Institute and the German Ministry for Health. And they pretty much pay the money and they pretty much tell us what to do, what they expect from the software, from the infrastructure and they are also responsible for the legal aspects, for example, where the app is released, which languages are released, and also they assume the responsibility. And of course, therefore some of the contributions we would probably have accepted in an open source, in standard community-driven open source

projects, couldn't be accepted in this special project just because it's basically a work defined by all partners. So that's basically the other side of the medal.

Karsten Hohage: How was the set up to reflect their role as a supervisory instance, basically in the community and in the overall operating model?

Sebastian Wolf: We had several workstreams set up on the side of Deutsche Telekom and SAP. For example, for open source, for development, for security, for data privacy and so on. And these had several meetings, sometimes daily, sometimes more often, sometimes also less often, with the partners on Robert Koch Institute's side and the Federal Ministry of Health. And we were, of course, part in these discussions. And on the other hand, we set up the normal open source process with a GitHub organization, repositories, we accepted issues, pull-requests and also featured requests on that end side. And while we were setting up these repositories, also the issued templates, so when somebody wanted to file an issue, file a pull request and a feature request, especially on the documentation side, we made it really clear that SAP and Deutsche Telekom were pretty much only the ones who are implementing that. Those responsible for the actual features, who are contacting us, are basically the Robert Koch Institute and the German Federal Government. Of course, we are publicly received as the ones who develop that. And that distinction is not always crystal clear. We had to explain that in a lot of cases. And of course, if something goes wrong, we are also being taken responsible that something is not working properly, though we probably might not be the culprit. Probably it might be in the infrastructure on the devices, it might be in the requirements, it might be because we implemented something wrong. In the end, we are in direct contact with the end user and we are the first contact when something is not really properly working. And we, of course, also needed to explain that in addition to, yes, we take it and we need to take care that it works again.

Karsten Hohage: Maybe let's go towards the grand finale with the two most interesting, positive, or critical feedbacks that you've got. I know about a very early one and about a very spectacular one. I'm dizzy. You might know what I'm getting at.

Sebastian Wolf: I'm curious, if I know that. So one of the most interesting things which I came across was probably the one interview with Linus Neumann from the Chaos Computer Club who was interviewed by Tagesschau. I think, if I recall correctly, asking

for feedback about the Corona-Warn-App from a data privacy perspective. And Germans or probably also Europeans know that the Chaos Computer Club is notorious for, let's say, being really, really harsh in that regards and checking it to the detail. And they said, well, or he said: "It's the first time, and it's also strange for me that I need to say this, they did it right".

Karsten Hohage: That is a very strong one from the CCC, the Chaos Computer Club. Just quickly, for the international listeners, Tagesschau which Sebastian mentioned, is like THE news show on German television. It has always been there and will always be there and everybody watches it. And the Chaos Computer Club, they go back to being day-one-hackers from back in the early 80s, right? Who probably hacked like the first instances of teletext and things like that and have now basically arrived in the IT mainstream society by being the experts for data privacy, security and so on.

Sebastian Wolf: Right.

Karsten Hohage: And then there was this early feedback about documentation that I remember.

Karsten Hohage: I heard this story about one of the very earliest feedbacks when you published the documentation was that it was not fully gender sensitive.

Sebastian Wolf: That was actually one of the first issues we received. I think on the first or second day when we published the first documentation. And you all know, it's a very delicate and sensitive topic. And of course, we as SAP and also Deutsche Telekom take care of that. And we haven't been that sensitive as we probably should have been at the very beginning. But that's also one important aspect. All these things, besides coding, that we need to take care of, we simply took the feedback, checked what's possible and what could be really extended when it comes to code of conduct and also ourselves, what we need to do when we write documentation and code, published a statement, and yeah, that was also really well-received in the end.

Karsten Hohage: What are your one, two or three main takeaways out of this Corona-Warn-App project, as a closing question?

Sebastian Wolf: The involvement of the public, also cooperation with public institutions, how to handle an open source project, especially if it's such a big one, properly and how to learn from it, how to do it better for the next project, why it's that way. Plan properly, prepare for the worst. That's really important. Be glad if the worst doesn't happen and also be open when it comes to feedback from the outside world. To incorporate this feedback in your daily work and try also to convince other people to, let's say, accept that there are many, many very good, positive, constructive elements and feedback out there that is just waiting to be incorporated in the product. And that brings us all further or makes us better. And as we said before, one hour invested to review the code, probably two or three hours in return. And that's also something I would take as a task for the future, to convince more and more people that this is really the essence of open source, that it's beneficial to invest in open source to get more back.

Karsten Hohage: All right, great. Thanks, Sebastian. That sounds like many people should remember that for as long as possible. Thank you, Sebastian, for being our guest here today.

Sebastian Wolf: Definitely appreciate it. Thanks for hosting me.

Karsten Hohage: Thanks for being here again. You out there, if you enjoyed this episode, please share it as often as you can and subscribe to the podcast in all the regular places. You may have encountered this on the SAP or openSAP website, but we will also be in all regular podcast channels. So spread the word and be back in two weeks when we'll have the next episode of The Open Source Way. Thank you.