

PPA Data for Impact Symposium

Discussion on Datastake with Benjamin Clair

Attendee 1: Could you elaborate more on the opportunity to monetize the data? And think about the expectation that supply chain due diligence data is shared down the chain, reconciling that with the aim to create data, and recognize its value. Is that a quality improvement issue? Could you delve into the dynamic, especially considering that a lot of the actors whose costs we're trying to amortize or share more equitably often don't have a lot of market power?

Benjamin: Thank you. First, we want to recognize the value of data on a financial level. As in, if a bank requires information on a mine site and requests it from a local NGO, technically that NGO could condition sharing of that data with the bank to a financial transaction. Although this is the long-term goal, the first thing we focus on is utility at application level.

By using the application that helps them organize their data, we try and help mining cooperatives meet, for example, regulatory requirements or requirements from their buyers in terms of supporting the due diligence process. For NGOs, we have built an application which they can use to streamline their work with a donor. Make it a little bit more efficient and attractive. I was mentioning this CSO user in Goma that got awarded an innovation prize from ILO, and that is because they have digitized their information on mine sites and on supply chain actors. They can showcase supply chain maps to various partners interested in the DRC. What that means is that you are funded autonomously as a locally based civil society organization, and can start demonstrating proper professional risk management. That's the utility aspect.

As for financial value of the data: OK, you've worked on risk management (because you were funded by USAID, because you were funded by ILO, or because as a cooperative you were asked to do so by your buyer), now you have all that data at your disposal and some other parties may be interested in what you have and may come to you. This party requests access to that information. And if you agree, maybe you can request them to pay a little extra for it. We don't anticipate it to be massive amounts, as the information may not always be complete, but certainly it's a good baseline before going to visit a mining area.

If you go on Google, you won't find much about mine sites, you can check the IPIS map, you can check reports on the area where you're going. On Datastake you could potentially see a bunch of users at local level and get their version of the local story (and also connect with relevant IPIS data). By making data super granular, we have transferred the capacity to label data to make it more useful. We transfer that capacity which platforms typically use to monetize themselves. We transfer it to users, making the data as useful as possible. We are seriously thinking about a more fluid market where you would go and query a location, a stakeholder, or an event, and identify all the holders of information in relation to that subject, and ask them for how much they're prepared to give you access to that data.

Attendee 1: It sounds like the data can be catered or curated for different audiences.

Benjamin: Yes. We try and be as objective as possible in the way questions are framed. Maybe I can show an example if I can share my screen. *[LIVE DEMO STARTS]* Let's see, so this is the screen of a local civil society organization.

I will show you how granular the data gets. I want to be clear that this is not too constrained by Internet access even in U.S. *[DISPLAY KYC SUMMARY]* So to get to this kind of an output on the part of this organization. If you go in 'Edit' mode, you will see that for every aspect of this cooperative there are granular data points. All the background, members, suppliers, clients, the level of scrutiny... we've built in a very detailed questionnaire with options for every question. Of course, they can add comments to expand on it. But the point is that all the options exist here. Actually, our users sometimes suggest new options to expand those repositories.

So that's the way this company or cooperative here is documented and explained by a local source. It's just facts reporting. It's all closed-ended questions (with, again, the opportunity to add the comments) to provide an analytical framework on what this cooperative is like. We have some appraisal sections, but they are not going to be consolidated into a client's own data. The same is true of my insights. Let me check a site where I know we have information... And by the way, I'm accessing this, and I should not. This is user data. Datastake has no rights over the data. But users understand that I may display some of their work at beta stage.

If I go on that mine site, it's the same thing. To describe that mine site, we have a very detailed process to identify risks at that mine site. I'll go in the risk topics that we develop the most, those that relate to child labor and modern slavery risk. If you go in there, you will see that the questions to characterize child labor at local level are extremely detailed – save for the victims' identities. We don't put names of victims in Datastake, obviously, because that data is sensitive and cannot necessarily be consolidated across multiple users. So it's less valuable from a statistical standpoint but, at an application level, some people are asking us already to put in victim tracking and things like that. Ethnicity, who are the perpetrators, type of victim, recruitment conditions, etc. We went to a great amount of detail on the way we present this. Same goes with the risk assessment methodology, proposing risk mitigation measures. As a local user, as a local civil society organization, you can organize your data for many different purposes.

For example, all the recommendations that this civil society actor is making can then be shared with the buyer. For example, at the site X, some stakeholder engagement would be needed. These are some details on the recommendation. As a buyer of gold from the site X, you receive all the recommendations made by local actors.

Attendee 2: How can we get in touch with you to get this application?

Benjamin: You can go to the Datastake website. There you can also register and even apply for the demo. If you want to enter dummy data and play around just check how the repository works. We also have a community forum where I just posted some updates, and you can join us on Discord. That's where you're going to engage with users actually and discuss opportunities and things like this. You can find a link to that at the bottom of the datastake.io webpage It's really a community forum.

Attendee 4: If we're working with a local NGO, are we able to have them connect to this and put in their data for the projects that they're completing as well?

Benjamin: Yes. Any organization is free to use Datastake, if they are interested to organize their data instead of just a report, an email. They can use Datastake to convey any information that they have on local mining communities and on villages, on supply chain actors, incidents. If that is the type of data structure that is of interest on your end, they can then create an information sharing project, as we call it, and invite you to access their information. When you create an information sharing project, it will be in the project section

here, you will create the project, and you will identify information clients and potentially other information sources.

And that newly created project will be accessible from this top section. In that project, you would access not only your own data but any data that those participating users have published. So if you have a local NGO that you can think of, which is working locally, they can try to process this right away. As soon as we set up the project back to see if they have data to compile on stakeholders on locations or events, because there's also an event reporting process here for incidents and corrective actions. The best thing to do would be to try it out to see if it's suitable and if it serves any purpose on your site.

Attendee 4: This seems like there's two kind of parts. My question is: obviously, as a downstream company, we rely on the OECD framework and smelters and refiners completing much of this due diligence that needs to be part of their process. Have you worked at all with any particular smelters and refiners that are located in the DRC here that are buying? Or is it also open to them to be able to use this data since they have more visibility on the mine sites from where they're buying?

Benjamin: We started with gold refiners that were taking part in the Zahabu Safi USAID project discussions in DRC. I previously set up a due diligence program in DRC. The situation of 3TG, 3Ts in particular is not really a field where I want to uh, to involve myself at the moment. So we haven't reengaged with the 3T smelters. But some gold refiners have expressed interest now. As comprehensive as the information sharing framework may be, there are other issues that have not been fixed in the DRC when it comes to gold procurement, the main one being that even if due diligence reporting system is free, it still makes more sense, for tax reasons and for other reasons, to channel your gold to towards illegal buyers. So even though some refiners were really interested in procuring from DRC gold sources, it's still pending. Not because of due diligence, but because they can't agree on price. They can't find a suitable price mechanism.

The platform just provides access to the data that may be needed for your due diligence or for your risk management or impact management. But then it's down to users to decide if they have enough data and other conditions are met for them to engage upstream and extend. End users should not offload the responsibility for local transparency to refiners and smelters.

Attendee 5: To what extent are artisanal miners involved in this data collection process?

Benjamin: The users of Datastake would be the cooperatives and the community-based organizations. Then those organizations are expected to compile information from community members or, in the future, integrate data that is already generated through offline mechanisms. We haven't built an interface for individual users, mostly because those applications exist already. Come and talk to our users on our [Discord community forum](#).

Datastake is an incentivized data management platform connecting information holders in countries of mineral origin with international buyers, investors, and partners. It implements the Consolidated Automated Due Diligence (CADD) framework to eliminate upstream compliance costs, facilitates locally-led monitoring of development interventions against the UN SDGs, and is used by banks to onboard ASM actors through triangulated KYC/AML checks.



Benjamin Clair is the founder of Datastake. He has 12 years of experience in the fields of risk management in CAHRAs and mineral supply chain due diligence, and previously created an upstream traceability and certification program which to this day remains the only legitimate solution to procure 3Ts from DRC and Rwanda. With Datastake, his objective is to transform information management in weak-governance countries, and financially incentivize transparent reporting.