

Center for Strategic and International Studies

TRANSCRIPT

Event

AI for Food Security Forum

“Technical Demonstrations – World Bank DIME AI”

DATE

Thursday, April 30, 2026 at 9:30 a.m. ET

FEATURING

Philipp Zimmer

ZeroHungerAI Project Lead, World Bank DIME AI

CSIS EXPERTS

Zane Swanson

Deputy Director, Global Food and Water Security Program, CSIS

Transcript By

Superior Transcriptions LLC

www.superiortranscriptions.com

Zane Swanson: Hi everyone. If you can please find your seats, we're going to get started with our first tech demo of the day. Thank you.

Hello, everyone. My name is Zane Swanson. I'm deputy director and fellow with the Global Food and Water Security Program here at CSIS. Thank you for being here, both in person and online. I hope everybody's been able to recharge and get some caffeine in our coffee break as we move forward with the first technical demonstration of the day. Very excited about this.

Join me in welcoming Philipp Zimmer ZeroHungerAI project lead at World Bank Group's DIME AI, as he presents a novel tool that anticipates outbreaks of food crises by analyzing news articles for keywords associated with drivers of food insecurity. It's a really great presentation. He's doing really important work. So join me in welcoming Philipp to the stage. Philipp. (Applause.)

Philipp Zimmer: All right. Good morning, everybody.

Now the magic question: AI demo, does the tech work? (Laughter.) It looks like we're good. Lovely. Yeah.

Thank you so much to the CSIS team for hosting this wonderful event, for having us here. Also, once again many thanks to the Google.org team for making this possible, also making our work possible that I will be speaking.

As Zane already so kindly introduced, I'm Philipp Zimmer. I'm an applied AI scientist at the World Bank Group's Development Impact AI Lab. And to give you a little bit of context, where do we sit? We sit within the operational arm of the research department of the Bank. And so we focus since about three to five years now our efforts on building context-specific AI solutions for some of the pressing development challenges, and where we also see the technology can very neatly fit into operational workflows at the Bank and outside.

Now, in our work on ZeroHungerAI, which is gratefully funded by Google.org, we are looking particularly into food insecurity forecasting. And to situate this area of work a little bit, I wanted to give an overview of, why do we actually care about this?

So, in practice, unfortunately at this day, a lot of humanitarian response is still very reactive. What does that mean? That we don't really have good mechanisms in place to provide early indications of known risk factors that drive through crises. That then leads to delayed response with significantly higher impact on loss of life and humanitarian cost.

Now, the idea behind anticipatory action and behind forecasting that informs thereof is that we can capture early warning signals across a variety of domains, thereby inform anticipatory action not necessarily to prevent crises, but to mitigate the effects thereof. And there's been a lot of research that has shown already the cost savings and can come from that, and also the savings in lives lost.

Now, as we started doing work on this – in this domain, one of the key motivators for the work you will see about in a second was that data availability still remains a key challenge when we're building food insecurity forecasting systems. And it goes as far as to say that compounding crises we observe today are oftentimes accelerating much faster than the systems we're building to track them. And this is very often also driven by the fact that traditional data that's being collected – whether that's from surveys, whether that is from administrative data from governments – unfortunately, these records and these signals are infrequent, incomplete, and very often – (audio break) – the promises of AI that we are observing, it still needs sustained investments in on-the-ground data collection. There is a lot of magic happening, but there is no magic without the data that can help us train these models and also evaluate them so we know whatever we deploy is rigorously tested and fit for purpose. And on the second hand – and this is what you will see in a second – we also want to think, though, about research innovations that can help us close these data gaps that I've outlined.

So a few years ago we started some work on the idea of, can we get proxy signal on these known risk factors we know drive food crises but we can't reliably observe at scale at high frequency? And the promise there was to test the hypothesis that news reporting, and particularly local news reporting, captures a lot of the risks that contribute to food crises much earlier than structured data does.

So in the end of the day, it's pretty simple. We collect large amounts of news content. We then track the discussions on known risk factors across different domains, from conflict, to agricultural events, to climate shocks, to economic crises. And then we turn these signal(s) that we are extracting into high-frequency features that can be fed into our prediction systems. And we see that we can do this pretty reliably and with pretty good coverage, up to the district (and two ?) level. And we then use this not only for the predictions, but also as a supplementary data source that policymakers can use. Because very often that's – what our user research has shown for our tool, people don't only care about the famous red dot on the map; they also care about understanding what is driving this crisis so their policy response can be tailored to that.

Now, beyond the promise, also some results. This is from a study we conducted a few years ago across 21 countries where there's reliable ground truth data. And what we've seen here, what we've attempted to showcase, it's the value of using a new signal to detect food crisis outbreaks, an outbreak here defined as the IPC level classification changing from below three to three or higher.

And the first good news is that traditional methods – and what we mean here with traditional methods is machine learning approaches that leverage some of the traditional data – already outperforms human experts. But something that's grounded a lot of the work that we are doing is that if you hold everything else constant – same model, still using traditional data – but then feed the additional signal that we can observe in the news, we have pretty significant spikes in our ability to identify crises.

Now, you might naturally be wondering, how does it come to be? And I brought two examples to showcase why this method is not only valuable for predictions, but also for interpretability of the forecasts so that we can actually give clear guidance to policymakers.

The first case is from South Sudan, where in the Mayendit District an IPC phase three was declared in 2017. And this crisis was driven largely by crop pests, and we saw this also observed in some of the traditional data that has been collected such as the vegetation index. It deteriorated significantly a few weeks before the crisis was declared. Interestingly enough, though, already months earlier we saw significant spikes in the relative mentions and discussions of pests and diseases in the Nubian region.

So essentially, what was happening was farmers were observing that new bugs they hadn't encountered so far were affecting their crops, and they were seriously concerned about how that would affect their harvests moving forward. And local journalists were essentially picking up on their chatter and were reporting on this in the news. So this is one example where news can help get an earlier signal.

Now, there's a second example I'm bringing from Somalia, where IPC phase four was declared in 2011, where this crisis was largely driven by conflict. What we found here is that some of the traditional indicators that are being used, such as fatality counts, were actually not making any official records, largely because the crisis was moving – the conflict was moving so fast. However, already four months earlier we saw again spikes in news reporting on violent events and conflict.

So those are just, you know, two examples of many of why do we think we can achieve these large predictive gains. It's not just magic; it's pretty interpretable, and we can bring it back to the signal.

Now, this is something I want to highlight, and maybe also piggybacking a little bit of one of Franck's closing comments. We're not here to reinvent the wheel, and we're not here to, you know, propose that this news monitoring can replace all these other amazing tools, but we really want to see it as a complementary source to a larger ecosystem that does a lot of amazing work across different domains trying to map and track different risk factors that play a role here. So we're very excited to be part of this research collaborative that's convened by Google where, you know, we're trying to be part of a mechanism and a group that really takes a bird's-eye view at this problem, bringing together all these different puzzle pieces and from our perspective supplementing that with rich signal across different domains with the news.

How do we turn this into practice? With the funding from Google.org, we've been building an open-access tool the last couple of months where users have access both to our predictions of a time horizon up to 12 months in advance and they're also supplemented information around the risk factors that are driving these crises. So, as you can see here on the right-hand side of the screen recording, you will not only be able to see where a crisis is going to occur, but also get some intel of what we observe in the news – where the chat is picking up on, on which topics we think is – are driving these crises.

So these interpretable signals come from more than 140 million articles across almost 100,000 sources, so we really try to take a broad look both local and global reporting. And this will be made accessible via the – an API, as well as the interface.

Now, the key question, I think, that also came up partly in the previous panel is, how do you operationalize this now, right? I started out the presentation thinking about the ways decisions are made, the ways in which forecasting models should feed into anticipatory action. So we're taking two concrete integration steps here, both within our organization with the crisis response window, which is a \$2 billion crisis response financing mechanism that's being set up together with national disaster management agencies of more than 30 countries; and then also in talks with the U.N.'s CERF, the response fund, to again provide this data as an additional signal, as an additional information source in the room for decision-makers,

Once again, thank you so much for the CSIS team for bringing us here. I would love to hear from all of you in the room any feedback, any harsh

critique that help us improve our tool, but also to see where there's complementary ideas of working together, you know, building this consortium, this collaborative in this space. Thank you. (Applause.)

Dr. Swanson: Thanks, Philipp.

We have just a couple minutes for some questions. So if anybody has a question in the audience, please raise your hand; a mic will come find you. Yes.

Audience member: Very interesting what you have developed.

I have a question about how do you distinguish the noise from the signal, correct? Because, yeah, you have all these datapoints, but how do you know where aid or whatever instrument is needed once you have so much information there? How do you make that evaluation? And do you cross this with information from the experts, but then the delay is there again? That's my basic question.

Mr. Zimmer: Zane, you want to take them one at a time, or?

Dr. Swanson: We can take one more if there's one more in the room. Right here.

Audience member: Yes. Karen Mancl from the Ohio State University.

What if you don't have trusted news sources, that the news sources – the news is generated by the government? Or what if you've lost a lot of the rural news outlets, as, you know, they maybe only were weekly and now they're – they have completely disappeared? Do you rely on social media? Or what are – you know, how else do you get the news?

Mr. Zimmer: Yeah. Both very excellent questions and questions we also think about quite a lot.

So, first, about the noise. So there's two elements that actually are pretty tricky that there are, though, pretty simple solutions for. One is, how do you actually account for deviances in reporting overall, right? There's not the same amount of news reporting everywhere in the world. But then also, two, how do you distinguish what is, you know, a baseline reporting on certain themes? And I think this speaks also partly to noise that's, you know, driving oscillations there.

So the first bit is although we see performance being pretty consistent across the globe irrespective of differing levels of news reporting, one of the things we do is we take – bring everything to the relative scale, so only because there's more reporting in Kenya than there is in Togo it

doesn't mean that there's more crisis in Kenya than there is in Togo. So when we speak about, say, discussions around conflict, we always think relative conversations to the overall news volume that we observe.

Then the second thing is that, obviously, there's different baseline behaviors. In certain countries, journalists might be focusing more on conflict. In other regions, they might be focusing more on weather events. So we're again normalizing our signal so that we capture, essentially, deviations from the norm, not so much what is a common behavior and the noisiness of news reporting.

So these are two of the mitigation methods.

Now, you point to a related point of, OK, what if there's a lot of crap going out in the world? Excuse my English. So all the modeling – and this ties very well with the – with the second question – that we do is validated with ground truth human labors. So all of the training we're doing with the models is not – you know, we're not validating against some baseline of what is good news reporting; you know, we are always validating our models with respect to what human experts have curated as, you know, ground truth.

And now speaking on deviations in news reporting and the prospects of social media, we have done some work with social media data in the past in a different context, but for two reasons that's not really appropriate. One, the problem of noise is even larger. And the second element is that the way platform restrictions currently operate it's not really a scalable method. So you can do this pretty decently for a country, maybe two or three countries at a time, but building a global system as we're aiming for is, unfortunately, not scalable.

Dr. Swanson: Great. Thank you so much, Phillip.

Mr. Zimmer: Thank you.

Dr. Swanson: Join me in thanking Philipp for his contributions. First tech demo of the day. (Applause.)

(END.)