URL: https://stvp.stanford.edu/clips/building-a-foundation-of-trust

Rahul Roy-Chowdhury, Grammarly's CEO, introduces the TRUE framework Grammarly has used while developing and launching their generative AI product, Grammarly Go. The first principle, Trust, is the foundation for everything, he says. He explains Grammarly's approach to making sure the company's incentives are aligned with their users when it comes to data collection and safety.

Transcript

- Everyone knows generative AI is the thing.. 00:00:05,250 Everyone's working on generative AI.. We've been working on generative AI at Grammarly.. And about a month ago we launched our take on generative AI, a product we call GrammarlyGO.. And so to bring GrammarlyGO to our users, we have millions of users around the world, we use the TRUE framework.. So let me walk you through it in detail.. How do we think about this framework when we thought about this product, that we just recently launched to our users? So let's start with trust.. Trust is about making sure that you understand what users' expectations are around their data.. These are questions like, well, what data are you collecting? These are questions like, well, what are you doing with the data? How long are you keeping it for? And really at the root of all of these questions, I think is an underlying question of, hey, what are your incentives? And are your incentives and my incentives, are they aligned or not? And so at Grammarly, we have built a business and a company around this idea of making sure that our incentives and our users' incentives are always in a hundred percent alignment.. And so it's a very simple decision-making process..

A new business model comes our way, and we think about it and we say, okay, there's revenue potential that's interesting, but does it cause us to lose alignment with our users' interests, and the way we are thinking about user data? And if the answer to that is, yes, we're not gonna do it.. So user alignment is critical.. That's really the root of privacy is understanding alignment.. If privacy is about alignment and what are you gonna do with user data, security is about keeping that data safe from harm, from misuse.. And at Grammarly, we've invested a lot in security over the years.. We have the whole alphabet soup of certifications, SOC 2, HIPAA, CCPA, you name it, we've got it.. And these things don't happen quickly.. They take a long time and take a lot of investment to do, and it's hard to bolt on after the fact.. And now with new technologies like large language models, new threat vectors emerge.. And the full landscape of threat models for something like a large language model isn't fully understood..

And so the best way to approach that uncertainty is to really have a multi-layered security story.. You're not just relying on one thing to work well.. You're relying on multiple things working together, having this portfolio approach, things like a bug bounty program, or things like having a red team.. The concept of a red team is you have a team, we have this team at Grammarly, and their job is to proactively find vulnerabilities in your systems so that we find them and fix them before an outside hacker does.. And so really investing in this portfolio of security approaches is the right way to deal with a new piece

of technology.. The thing I just want to kind of give you as a takeaway for this whole trust idea is it is very hard to earn, it's very easy to lose, and it's very hard to bolt on after the fact.. So as you embark on your careers, and you think about all the cool apps and businesses, and technologies you're gonna build, keep trust in mind, because trust needs to be the foundation of everything else you do because we need to earn, and strengthen that trust as the foundation.. It's gonna be very hard to scale a successful company without that basic foundation in place.. So that's trust.. So with GrammarlyGO, we applied this framework, these technologies, all of our investments..

They helped us launch GrammarlyGO, keeping our trust posture intact with our users...