BUZZ CHRONICLES > TECH
Saved by @ThomassRichards
See On Twitter

Twitter Thread by Lea Kissner

Lea Kissner
@LeaKissner

Next up in Privacy Technology at #enigma2021, Kelly Huang from @ethyca speaking about "GONE, BUT NOT "FORGOTTEN"—TECHNICAL & PRACTICAL CHALLENGES IN OPERATIONALIZING MODERN PRIVACY

Just imagine there's a global pandemic forcing everyone to stay home and buy their stuff over the internet. And you've been working on your sanitization-on-demand startup. You've got more users than you can count! ... literally, because your data's all over.

Now you're a multi-national international country with privacy issues because your information is all over the place.

Now a user writes to request you delete their data. Where is it? How do you do that? Who's responsible for privacy in your business.

How do you operationalize privacy rights?

Primary stakeholders:
* Legal
* Business
* Engineering

We spend a lot of time on Twitter analyzing the legal rulings, but it's harder where the "rubber meets the code" ■

Three rights:
* access
* rectification
* deletion

Legal's trying to uphold them, but it's a technical question!

Legal wants to decrease risk but don't know software

The business wants to stay in business and make money. They want to be able to use data for things like placing ads and analysis.

It takes a lot of time to handle these requests, too!

They need a streamlined technical solution.

Program management wants to streamline and make things efficient and predictable... but they don't understand the technical limitation

As a software engineer, you've seen technical debt. So much technical debt. All the weight of the decisions that were made in the past, especially if you scaled without a data plan.

Average SMB has data in 10 different systems.

How do we delete?
Some poor software engineer is trying to track down what data is where?
What even *is* PII? There's no real standard.
What should be returned? What should be deleted.

Make a definition and stick to it.

Some of your databases might use email addresses as a primary key, some user IDs, etc.

1. Define PII
2. Find all the PII
3. Use pseudonymization to replace PII with some kind of random value which can't be tied back to the user

[reminder I am livetweeting this is not me speaking]

How do you do this at scale?
Maybe a centralized team who can handle this?
If you're a small company, plan ahead!

Be careful when you're doing sanitization -- some databases really don't like batch processes and you can make things fall over.

Speed
* you have a timeline -- often 30 or 45 days
* but that's not enough time if you haven't planned for streamlined speed

Ideally you won't need it, but have a backup plan, in case something goes wrong with a slow data system

Plan for a solution that grows with your business, not just a hacked-together series of SQL queries, but instead a centralized portal with extensibility as the business changes and technical systems grow.

... and as new privacy laws come into place

Privacy is way, way more than compliance. But compliance needs to happen.

Let's all do our part

[end of talk]