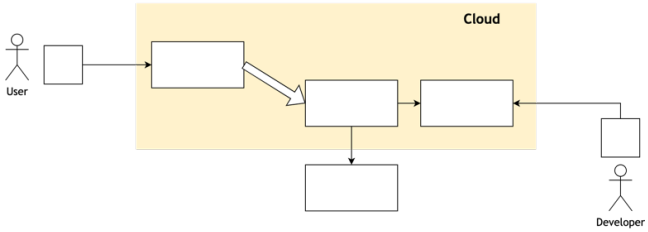
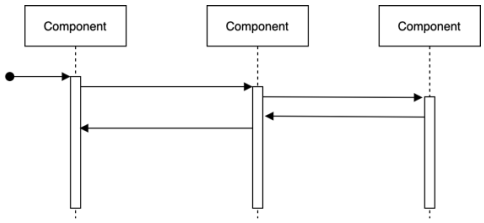
Software Diagrams
For Privacy Professionals

Attending a threat modeling or design review session? Here's how to look at diagrams with privacy glasses on (and make annoyingly concrete questions).

In most cases, you're going to see either

Data Flow Diagrams (DFD) **and/or** **Message Sequence Charts (MSC)**





In either case, both diagrams can be "interrogated" to effectively discover privacy risks.

Completeness and trutfulness
Does the diagram show *all* the data flows that contain personal data?
Is the personal data content of any of the data flows unclear or undecided?
Do any data flows end up in a void? Where does that data go, if it's not drawn?
If data storage has not been drawn, where is it located? It must be somewhere.
Do some data flows pass through another system not pictured – like the user's browser?

Data ownership and access
Are some of the shown components owned by other teams, orgs, companies?
Is there a contractual or geopolitical boundary between some of the components?
Where do developers (or admins) technically have access to? And from where?

Data lifecycle, minimization and inference
Does a component get personal data from several sources? What is the combination like?
What's the retention policy of each data store? How is it technically enforced?
For each data flow out of a component, does it contain personal data that is not needed?

Useful thought exercises: What if…?
Would it surprise the user if they saw these data flows?
Look at the components "on the edge". Assume they're hacked. What would be at risk?

By Antti Vähä-Sipilä, avs@iki.fi. https://fokkusu.fi/sdfpp.pdf

Version 2023-03-01 (IAPP DPI23)