EDM Webinar

Can Privacy & Data Sharing Coexist?

Live Date: April 30, 2020

Featuring: Scott Albin, Head of Ecosystem Services, Data Republic

Recording: bit.ly/2WsTypU

Presentation: bit.ly/35Ttw1R

edmcouncil.org   **datarepublic.com**

WEBINAR Q&A:

You spoke quite a lot about customer data. Is that the purpose of data sharing or does it extend more broadly? How so?

Data sharing extends broadly across all sensitivity levels of data from open data, transactional data, customer data, health data and more. We focus on customer data where data innovation is the strongest and the business cases are most compelling for many B2C organisations. Any organisation, public/private, B2C or B2B has data assets that are valuable to others and vice versa. The opportunity to expose this data, safely and securely, is enormous. We are here to enable all forms of data collaboration.

In the control section you discussed "use", how do you enforce these rules in ecosystems. Do you monitor? Do you require evidence from the user?

"Use" in the context of consent refers to the organisation taking action of some form, such as direct marketing, based on the data collected. We enforce standards of practice through our Common Legal Framework and also through the extraction controls and audit controls we have in place. The Common Legal Framework requires participation from all parties involved and provides various protections. Each project also has a data license that is agreed to by all parties within the project. The license covers the use of the data and the outputs of the project. We do not actively monitor what happens to data that passes through and out of the Data Republic platform.

As privacy laws are changing are the existing ecosystems able to adapt, are they grandfathered or are some dissolving?

Typically, we do not see grandfathering of privacy laws. They tend to evolve slowly and often in consultation with industry, but practices vary from market to market. We have designed Data Republic's platform in such a way that we can be flexible to changes in regulation, e.g. approved privacy preservation techniques and data sovereignty requirements.

Can you discuss scale issues?

Typically, what limits organisations from scaling their data collaboration efforts is the overhead of legal, risk, data protection, technology and other dependencies required to "approve" a project. While these functions are critically important, they often erode business cases and require significant manual effort. Data Republic offers standardised governance workflows, technology and a common legal framework, which allows organisations to scale effectively without increasing headcount. One organisation went from <5 projects per year to >25 per year with the same number of dedicated staff after deploying Data Republic.

Can you provide specific examples of Matched-in and Matched-out data use scenarios for the Financial services sector?

Matched-in and Matched-out refer to the flow of data. For example, matched-in refers to customer data, such as banking transactional data for an individual customer, being uploaded into the Data Republic platform and "matched" with a partner organisation's customers, such as a bancassurance insurer. Matched-out is the same transactional data, potentially transformed into propensity scores or segment identifiers, being extracted out at the individual level to the insurers data systems for use.

Could you elaborate on what you mean by more sophisticated with consent? Do you mean it is to focus consent on FIPPs for future legislation or maybe is it updating consents to be more precise when new regulations are passed?

Today consent is largely handled by large, onerous privacy policies presented to customers during onboarding and occasionally updated. Technology companies update these frequently. We believe in a future where consent is more dynamically managed down to the data attribute level. This will enable nimble collection of consent for use of that data, unbundling it from other data, allowing it to be time-bound and presenting more flexibility to end customers to control it. This future will begin to emerge over the next five years and organisations will scramble to catch up and seize the advantages from it.

Data assets are key facts for business, in terms of policy, privacy - Is data asset a must to have to bind with all the internal vertical of organisation data?

Data from any business function has a certain level of utility, either to the data custodian who produces and "owns" it or to others in the same business or potentially with business partners externally.

Data Assets that drive the business values must be adopted by all organisation. What's your thought?

Correct. At Data Republic we are enabling organisations to extract the value from this data, which we refer to as the "utility of the data" either within their organisation across silos or externally.

How does your technology work? SaaS remote, on-prem servers, etc

Currently the Data Republic platform is SaaS based with an on-prem "Contributor Node" for privacy preservation and secure transmission to our quarantined analytical workspaces, which are on AWS. We are progressing our product roadmap to support multi-cloud and on-prem deployments.

Please explain the "pink Ferrari" possibility expression. How can the "pink Ferraris" be identified in a data set?

This is referred to as re-identification risk. It is possible even with identifiable data fields such as name removed to re-identify someone in a dataset from the other data attributes. For example, in a record of car sales with no identifiable data included, there may only be a single "pink Ferrari" recorded, which could help in identifying the individual associated with this data.

Some companies want to save all data because they don't know of any future use case yet. However, to reduce risk compliance groups don't want to save data. What are your thoughts on how to balance this?

Typically, we find that more recent data is more valuable. For AI application there is also a limit to the usefulness of large data sets, where there are diminishing returns in terms of performance for adding more data. As a result, we do not see a significant issue with old data being erased over time. Further the relevant signals from this data can be chosen to be retained in more aggregated form.

Do the Data Republic products, where Privacy is important, also enable focus on confidentiality e.g., AICPA requirements?

Yes, our platform is used to host sensitive data where organisations do not want to lose control, but see value in provisioning access to certain individuals or organisations temporarily.

I'm interested in exploring data sharing further. What steps should we start taking now to prepare?

At Data Republic we always advocate for a strong focus on the use cases, business case and clarity of purpose on who you want to engage in any form of data collaboration. If you want to focus on innovation with your data assets, our Innovation Sandbox solution will help you connect your data with the outside world of experts, startups, and academics fast. If you want to merge data with a partner organisation, our Data Collaboration Suite will provide you the governance and technology required to do so securely.