GDPR Perspectives:
Privacy Risk Automation for GDPR

Your Customers’ Privacy, Protected!
www.bigid.com • info@bigid.com • @bigidsecure
GDPR Perspectives: Privacy Risk Automation for GDPR

One of the EU GDPR's regulatory directives is a new emphasis on risk-based privacy protection. However, the GDPR offers no clear definition of what constitutes privacy risk. The Regulation only pinpoints what it considers the riskiest processing – large-scale processing with the intent of profiling – and how to lower risk through more data protection, including pseudonymization. Determining where data processing lies on the risk spectrum, requires not only cross-functional collaboration and a responsive risk model – it also requires a clear view of all the personal and private data and access to that data that an organization is responsible for. This whitepaper looks at quantitative, data-driven software approaches to measuring privacy risk and using it to enhance data protection and privacy compliance.

Introduction

Along with the European Union General Data Protection Regulation (GDPR)'s stress on data subject rights – where consumers will have legal rights and access to their data long after it's collected by a company – the Regulation's emphasis on minimizing privacy risk will require a rethink of how personal data is handled by companies operating in the EU.

So central is the notion of risk to the structure of the Regulation, that the word appears 75 times in the current version. Further reinforcing the point, risk assessment appears in the very first sentence outlining the responsibility of the controller (the entity that determines why and how personal and private data should be collected):

"Taking into account the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for the rights and freedoms of natural persons, the controller shall implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with this Regulation."
The promulgation of a risk-based privacy framework is a monumental departure from traditional compliance requirements. Rather than only require evidence of controls and logging of activity, the GDPR goes further by requiring controllers to also think about the probability of a failure in controls and safeguards, and then gauge the harm resulting from a failure to an individual's privacy. Those obligations are encapsulated in the Data Privacy Impact Assessment provision that mandates covered organizations proactively assess the risk to private data involved with collecting and processing and whether adequate controls and data security protections are in place – even before a new service or application is launched.

The question that immediately faces any organization that will be subject to the EU GDPR is how to adequately and systematically measure privacy risk. Risk, after all, is not at first blush easy to quantify or take action on.

Adopting technologies that help businesses know where their personally identifiable information is, and who that personal data belongs to, and who is accessing that data are critical questions that must be consistently answered for any sustainable approach to risk analysis.

This white paper will review how the GDPR incorporates risk into its requirements and how organizations can leverage risk-oriented privacy management software like BigID to both identify data and usage risk and then automate actions to reduce that risk.
The Risky Business of Risk Assessment

Despite risk assessment figuring so prominently in the GDPR, the definition of what constitutes risk and how to analyze it is not explicitly stated. The Regulation affirms that the “likelihood and severity” of potential harm to personal privacy should frame the assessment of risk, but it doesn’t provide explicit technical guidelines on how best to reduce the likelihood and severity of exposure of personal and private data.

While the GDPR is not explicit in how to quantify risk, it provides a number of hints. Since the GDPR states that covered organizations should take risk into account as part of their “general obligations”, it’s reasonable to assume that risk analysis should extend beyond how the data is protected to how data is processed and whether those processes are in line with stated policies and user consent agreements.

In that context, any activities that involve large-scale processing and evaluation of personal data characteristics for the purposes of profiling elevate the potential for risk of harm. The GDPR defines harm very broadly as “physical, material or moral damage”, from processing activities that could lead to “discrimination, identity theft or fraud, financial loss, damage to the reputation, loss of confidentiality of data protected by professional secrecy, unauthorized reversal of pseudonymization, or any other significant economic or social disadvantage”.

The risk of harm posed by processing should in turn ensure that covered organizations are implementing “a level of security appropriate to the risk”, according to Article 30. For instance, GDPR considers processing to be low risk when it involves pseudonymized data and is not intended to yield personalization results.

Quantifying risk and assessing the probability of harm, therefore, needs to begin with:

- Understanding what and whose data is being collected,
- Measuring the sensitivity of the data attributes
- Assigning the residency of the data
• Mapping how it is being processed and accessed
• Pinpointing where and how it is being stored
• Identifying where it is being misused, either on the basis of processing violations or unauthorized access
• Analyzing whether the data is being collected in line with consent agreements
• Analyzing whether the data being collected is aligned with purpose limitation
• Determining when it is vulnerable to exposure or theft.

**Data-driven Risk Modeling**

Managing risk and limiting the potential of harm to data subjects is, of necessity, the outcome of the sum of the parts - privacy policies, data collection processes, application design, storage security, data management, access controls and data protection. Different stakeholders, whether privacy officers, IT departments, application owners or information security professionals, will come at the question of risk from a range of perspectives. Technology should serve as tool to understand, quantify and manage privacy risk for each of these functional areas.

Insight and intelligence into data flows, access activity and data mapping are the foundational inputs needed to move toward a cross-functional risk management and mitigation framework. In order to provide actionable recommendations on high-risk activities and drive automation of workflows - whether by data management or information security teams - risk models must be fueled by real data.

BigID’s risk modeling tool helps organizations understand and compare data risk based on factors that are central to the GDPR’s requirements including data sensitivity, residency, data security and application access, but presented in a view that is relevant to the stakeholder.

Factoring in both static inputs, notably data residency and relative sensitivity, and dynamic risk inputs, like analysis of access patterns, application access, but presented in a view that is relevant to the stakeholder.
Factoring in both static inputs, notably data residency and relative sensitivity, and dynamic risk inputs, like analysis of access patterns, application behavior, authorization intelligence and data vulnerability produces a much more comprehensive picture than a series of disjointed and disconnected (often manually intensive) processes conducted in organizational silos.

BigID’s risk assessment and analysis is designed to be practical and immediately useful as a KPI measurement tool. BigID provides an operator a simple configurator to set weightings based on an extensible pallet of risk inputs like data type, data residency, consuming applications etc. Administrators can define specific weightings to define a specific risk KPI. Moreover administrators can generate any number of independent KPIs so they can track KPIs against any number of benchmarks similar to how stocks can be benchmarked against multiple index performance benchmarks (Dow 20, S&P 500, ..).

Building a risk model on this basis, also allows stakeholders to drill into specific data or data activity throughs simple filters or queries thus making it easy to assess risk on any slice of the data or data usage. Also, risk mitigation recommendations can be ordered by their relative impact on the total risk.

BigID provides a set of visualization tools and role-specific views that helps all stakeholders understand risk performance over time: whether an individual application owner with remit for a single application, or the chief privacy officer with responsibility for privacy protection across the organization. Moreover any of the risk metrics can be accessed via the BigID API making it easy to re-use the risk KPI in a GRC tool, a SIEM analysis tool or even a UBA security tool.

These automated risk assessments can also be used in conjunction with BigID’s data discovery and data mapping tool to prioritize investigations or assess high risk processes requiring Privacy Impact Assessments – a requirement of GDPR. Based on the risk intelligence provided by BigID around data and usage, privacy officers can prioritize risk mitigation efforts.
No Two Risk Profiles are The Same

Because every organization will have a different emphasis and risk philosophy, the BigID risk model can be customized to allow organizations to set the weights and thresholds that are relevant to their risk profile. For example, an organization may choose to give higher a risk profile to individuals residing in the EU relative to their other data subjects, while others would assign a higher risk profile to data processing that takes place outside of the EU.

In verticals such as online retail for example, companies will assign a higher risk to associating multiple identifying attributes to what should be anonymous Web browsing data. The BigID system can be configured to detect when data captured from cookies is being joined with personally identifiable information, or even whether a processing partner is accessing both data stores - and generate an alert with a set of actionable recommendations.

In the medical research sphere, the relative risk weighting given to all personal data attributes would be higher still, and the risk model would be weighted more on the probability of re-identification through processing. In this instance, the BigID system can be used first to inventory PII data, identify where pseudonymization should be in place, and detect when a processor or application is accessing data with comparable attributes, but no data protection is in place.

Since risk modeling is still an emerging area, BigID's risk assessment
Conclusion

Even as uncertainty prevails on how to best align with the EU GDPR’s risk-based framework, it’s clear that covered organizations will have to grapple effectively with both quantifying risk and cross-functional processes for mitigating risk. Also, while it’s clear what will be considered high risk activity, and that risk can be lowered through more effective protection, risk is not static and there is no one size fits all model – either for organizations as a whole or for the stakeholders that are now all accountable for managing risk.

BigID provides a simple, automated way for organizations to define one or more data and data usage risk models and then monitor their data and data usage against the corresponding KPIs.