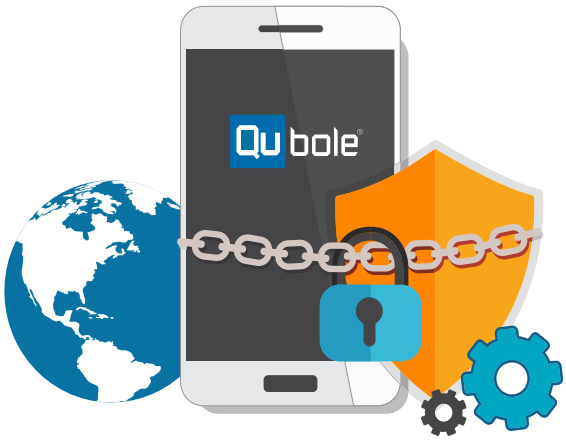
Qubole®

GDPR:
Security & Compliance

WHITEPAPER

Overview



No other company can match Qubole's prowess at the intersection of Big Data and the cloud. Qubole was founded by real-world operators who understand that security, confidentiality, and data privacy are fundamental to our mission and our commitment to a customer-first culture. We understand that GDPR compliance may be a important thing for you and Qubole is prepared to support GDPR and your compliance and regulatory needs. Qubole is committed to using our domain knowledge and best practices to help you meet the GDPR regulations.

In this whitepaper, we discuss:

- Qubole and the GDPR regulations
- Our compliance strategy—how we prepared for the GDPR from the start
- The shared security model—your responsibilities as a Qubole user
- How Qubole can help with GDPR compliance

What is the GDPR?

The EU General Data Protection Regulation ("GDPR") is a comprehensive data protection law that updates and replaces the Data Protection Directive 95/46/EC for all EU member states and is designed to strengthen the protection of "Personal Data" (any information relating to an identified or identifiable natural person, so called "data subjects") in light of rapid technological developments, the increasingly global nature of business and more complex international flows of personal data. GDPR will be directly enforceable in each EU member state. **The GDPR takes effect on May 25, 2018.**

What Constitutes Personal Data?

The EU defines "Personal Data" as "*any information relating to an individual, whether it relates to his or her private, professional, or public life. It can be anything from a name, a photo, an email address, bank details, posts on social networking websites, medical information, or a computer's IP address.*" The new obligations pertain to any organization that handles data about EU citizens—whether that organization is in the EU or not. The regulation does not apply to the processing of personal data for national security activities or law enforcement.

Does My Organization Need to Be GDPR Compliant?

If you are processing personal data within the EU, the GDPR applies to your organization. The GDPR also applies if you are processing personal data on EU subjects but your organization is not located in the EU. In other words, the GDPR is specific to where the person whose data is obtained is located or where the data was collected—not where the processor is located. "Processing" means any operation performed on personal data, such as use, storage, analysis, aggregation, transfer, dissemination or erasure.

What if My Business Isn't Located in the EU?

If your business is not located in the EU, the GDPR applies to you if you are offering goods or services (whether paid or free) to EU data subjects (data subjects are defined as EU citizens or EU residents) or monitoring the behavior of EU data subjects within the EU. Monitoring can be anything from putting cookies on a website to tracking the browsing behavior of data subjects to high-tech surveillance activities.

Note: Unless your organization can track the origins of your data including the time that the data was tracked and the dates of collection, it is possible that all your data could fall under the purview of the GDPR.

Controllers and Processors

Under the new GDPR legislation, organizations processing personal data are divided into "Controllers," or the entities which control the personal data, and "Processors," the entities that process personal data only on the instructions of the Controllers. The GDPR applies to both Controllers and Processors.

KEY REQUIREMENTS OF THE GDPR

Although the GDPR is daunting in its complexity and scope (there are 99 articles in total), the key requirements governing data collection processes (commonly referred to as the Seven Key Principles) are summarized below:

1. **Lawful, fair and transparent processing**–emphasizing transparency for data subjects.
2. **Purpose limitation**–having a lawful and legitimate purpose for processing the information in the first place.
3. **Data minimization**–making sure data is adequate, relevant, and limited, and organizations are sufficiently capturing the minimum amount of data needed to fulfill the specified purpose.
4. **Accurate and up-to-date processing**–requiring data controllers to make sure information remains accurate, valid and fit for purpose.
5. **Limitation of storage in a form that permits identification**–discouraging unnecessary data redundancy and replication.
6. **Confidential and secure**–protecting the integrity and privacy of data by making sure its secure (which extends to IT systems, paper records, and physical security).
7. **Accountability and liability**–demonstrating compliance.

How Qubole is Preparing for the GDPR

Qubole complies with the GDPR in the delivery of our service to customers and is fully prepared to handle the intricacies of the GDPR legislation. Specifically, Qubole will also continue to enhance data protection and compliance in the areas below.

Accountability

Data Protection Officer (DPO)-Qubole employs a Chief Security Officer (CSO) who functions as the organization's top executive responsible for security. Our CSO will also serve as our DPO under the new GDPR requirements.

Policies and Procedures

Qubole maintains a set of security policies, standards and procedures that provide our workforce with stringent data protection and compliance guidelines.

Qubole Data Protection Policy Addendum and Agreement-Qubole has created a Data Processing Addendum as an attachment to its Master Services Agreement. This document supports our commitment to this important legislation and is available here.

Mandatory GDPR Security Awareness Training

All Quboler's have taken mandatory GDPR compliance training. Further job-specific training will be required for individuals with responsibilities related to GDPR compliance.

Compliance and Risk Activities

Qubole evaluates the design and operation of the Qubole platform, including all services, applications and processes to ensure compliance with internal and external standards. We engage credentialed assessors to perform external audits at least once per year including TrustArc.

Privacy Shield-Qubole engaged with TrustArc (formerly TRUSTe) to complete and attest to compliance with the US Privacy Shield regulation around privacy and transfer of EU Personal Data to the United States and now works with them for arbitration and notification services.

3rd Party Innovation

Qubole has partnered with 3rd party GDPR specialists to provide translation and assistance with the practical application of the GDPR to our unique business model.

Data Protection and Other Security Measures

Customer Data Protection-GDPR regulations mandate that personal data is kept confidential and secure. Qubole's maintains a staff of security practitioners dedicated to ensuring that all systems (IT, and Development) remain secure and confidential at all times. For detailed information about our security program, please refer to our whitepaper, *Qubole on Amazon AWS: Security and Compliance Whitepaper.*

Design Security-Qubole assesses the security risk of each software development project according to our Secure Development Lifecycle. Before we complete the design phase, we do an assessment to qualify the security risk of the software changes introduced. This risk analysis leverages both the OWASP Top 10 (discussed below) and the experience of Qubole's Product Security team.

3rd-Party Suppliers-To run its business efficiently, Qubole relies on a limited set of sub-service providers. In areas where those sub-service providers could impact data security, Qubole ensures that service organizations adhere to confidentiality commitments Qubole has made to its users. Additionally, Qubole monitors and reviews all sub-service security safeguards by conducting reviews of its service organization controls before use and at least annually.

Data Breach Notification and Reporting Requirements

Notification and Customer Communication-Qubole makes every effort to maintain the security of customer data. In the event that an incident occurs that exposes or provides unauthorized access to data, Qubole will respond to any impacted customers no later than 72 hours after the event.

Any sensitive or confidential information will only be shared with authorized users.

What are My GDPR Responsibilities as a QDS user?

Similar to our existing legal requirements, GDPR compliance requires a partnership between Qubole and our customers in their use of our services. Security in the cloud is slightly different from security in your on-premise data centers. When you move computer systems and data to the cloud, security responsibilities become shared between you and your cloud service provider.

Each party-the cloud provider and cloud user-is accountable for different aspects of security and must work together to ensure full coverage. When you use a cloud provider, they are responsible for securing the underlying infrastructure that supports the cloud, and you are responsible for anything you put on the cloud or connect to the cloud. This model of shared security responsibilities also applies to Qubole as a service provider and your role as a QDS user.

QDS and Your Data

One of the most unique facets of QDS is that even though the service is provided in the cloud, Qubole does not need direct access to your data. QDS is architected as a service platform with three primary components:

- The big data applications (Spark, Hadoop, etc.) along with additional components to help leverage these technologies such as Hive, Pig and Tez and, finally, the storage layer of HDFS and/or file-based storage on Amazon S3.
- The orchestration infrastructure that takes desired inputs from the customer (whether to use on-demand or spot instances, the minimum and maximum size of the clusters, whether or not to encrypt, and dozens of other inputs).
- Finally, the interface itself, translates the complex command structures of big data by providing an easy mechanism for customers to create, test and run their queries and various commands.

Role-Based Access Control in AWS, Microsoft Azure, and Oracle Cloud

Qubole also uses Identity and Access Management (IAM) roles in AWS and other types of Role-Based Access (RBAC) in Azure and Oracle Cloud to limit access to resources such as storage and compute by using a refined set of permissions. This allows our customers to use Qubole on their behalf by granting limited access to process the data in your cloud provider account.

Additionally, common concerns are addressed including limiting access rights to modify or affect the status of clusters, limiting the types of commands your users can execute and the data engines they can use.

For more information on IAM roles, see Managing Roles in QDS and our technical paper *Authorizing AWS in QDS—Using Secure AWS IAM Roles and Policies.*

The following table illustrates the respective shared security responsibilities between your organization, Qubole, and your cloud provider with respect to GDPR compliance. (To better understand your cloud security responsibilities, please refer to your Qubole Service-level Agreement).

Customer	Qubole	Cloud Provider (AWS, Azure, Oracle Cloud)
Responsible for user access management and data security in the cloud	Responsible for security of the platform and big data service in the cloud	Responsible for the security of the cloud
Customers own the data and are responsible for the security of their data Qubole is granted rights to process the data in the customer's account	Responsible for secure access to data platform	Responsible for secure storage in the cloud
Responsible for data encryption	Responsible for secure transport of commands	Responsible for availability and redundancy in the cloud
Responsible for user management	Responsible for multi-factor authentication for administrative access to systems with more highly sensitive and regulated data	Responsible for compute resources in the cloud
Responsible for infrastructure identity and access management	Responsible for operating system, firewall configuration	Responsible for networking in the cloud
Responsible for Qubole groups and role definitions	Responsible for metadata security	Responsible for encryption technology, key management capabilities
Responsible for data residency Responsible for requesting and reviewing 3rd party attestation and certification reports	Responsible for 3rd party attestation/validation (SOC2, HIPAA, PCI)	Responsible for 3rd party attestation/validation (SOC2, HIPAA, PCI)

Leveraging Qubole for GDPR

Qubole provides the following functionality that can assist you with data governance and security.

GDPR requires you to:	QDS allows you to:
Control Access to Personal Data A pillar of GDPR is limiting who has access to crucial data in your domain. While it sounds simple, consolidating a list of administrators is tricky. Limiting and tracking access to your applications can prove even trickier.	**Enforce a Least Privilege Model** You can restrict who can view, create, edit, and delete your most sensitive data objects using: 1. Privileges based on job requirements using Identity and Access Management (IAM) roles in AWS and other types of Role-Based Access (RBAC) in Azure and Oracle Cloud. 2. Qubole also supports Amazon AWS Key Management Service (KMS). This service stores keys in a shared Hardware Security Module (HSM) dedicated to encryption key storage. 3. Qubole provides a policy document that is GDPR compliant. This policy defines the permissions that apply to a user, group, or role; the permissions in turn determine what users can do in AWS.
Follow the Right to Be Forgotten This rule allows a person to request that any data a company owns about them be deleted. This can be anything about a specific individual, ranging from a social security number to a CRM record. While this rule cannot supersede another law (like a requirement to maintain HIPAA records), it is essential for any company who houses personal data online.	**Discover and Delete Data** Qubole has a data deletion process that allows you to comply with requests from individuals exercising their right to erasure. **Note:** This process only covers the Qubole customer data that is collected in the QDS service and not the customer-managed data within their environment. Please contact your account representative for access to Qubole's written data deletion process.
Report a Breach in 72 Hours In the case of a personal data breach, the Controller needs to notify their local Data Protection Authority figure within 72 hours after becoming aware of it. Companies should have a cross-functional incident response plan prepared that includes the Public Relations, Legal, Compliance, IT, and Security teams.	**Defined Policies and Procedures for Incident Reporting** Qubole has strictly enforced security policies that govern all aspects of our incident reporting process. Please contact your account representative for access to Qubole's written incident reporting process.

Conclusion

Qubole is prepared to meet the challenges of the new GDPR legislation and we want our customers to know that we take data protection seriously. Qubole understands that GDPR compliance is a shared effort between our organization and our customers. It requires a combination of people, process and tools and to that end, we are committed to helping you prepare for the GDPR. We welcome your questions and would be happy to discuss the ways that we can work together to ensure GDPR compliance for your organization.

ABOUT QUBOLE

Qubole is revolutionizing the way companies activate their data—the process of putting data into active use across their organizations. With Qubole's cloud-native Big Data Activation Platform, companies exponentially activate petabytes of data faster, for everyone and any use case, while continuously lowering costs. Qubole overcomes the challenges of expanding users, use cases, and variety and volume of data while constrained by limited budgets and a global shortage of big data skills. Qubole's intelligent automation and self-service supercharge productivity, while workload-aware auto-scaling and real-time spot buying drive down compute costs dramatically. Qubole offers the only platform that delivers freedom of choice, eliminating legacy lock in—use any engine, any tool, and any cloud to match your company's needs. Qubole investors include CRV, Harmony Partners, IVP, Lightspeed Venture Partners, Norwest Venture Partners, and Singtel Innov8. For more information visit www.qubole.com.

Qubole®