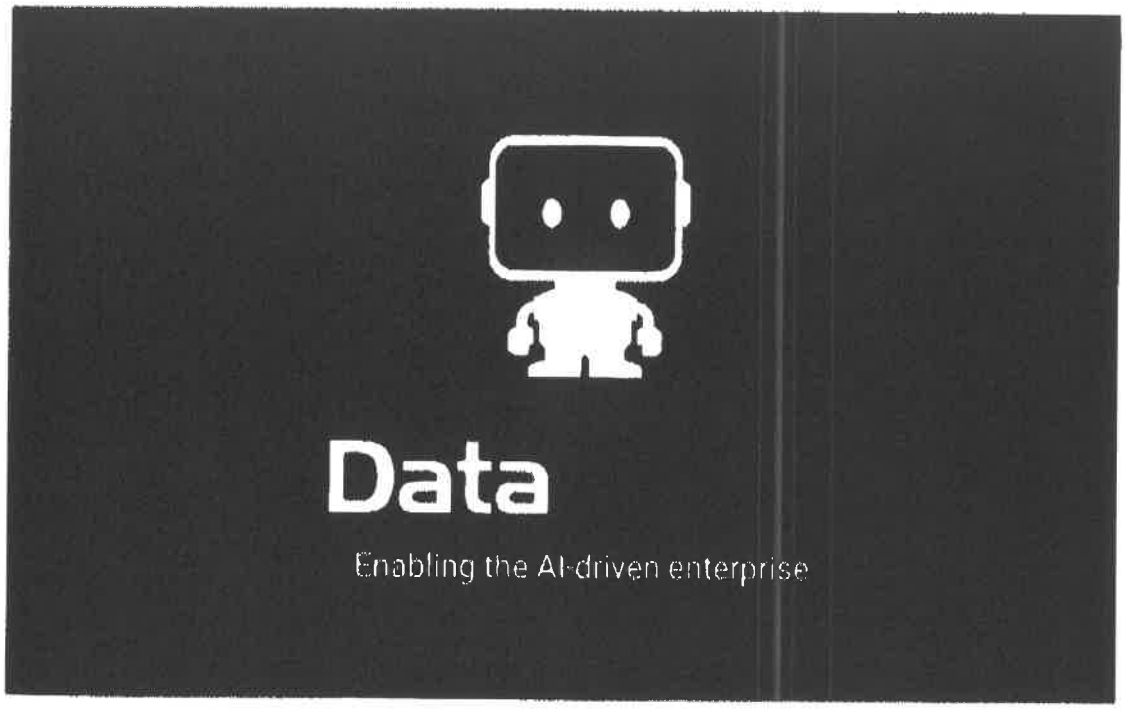
DataRobot Technical Proposal COVER PAGE

Bid Delivery Address and Fax Number:
Department of Administration, Purchasing Division
2019 Washington Street East
Charleston, WV 25305-0130
Fax: 304-558-3970

September 28, 2022

Vendor Name:	DataRobot
Buyer:	Crystal Hustead
Solicitation No.:	CRFP MIS2300000001
BID Opening Date:	September 28, 2022
BID Opening Time:	1:30 PM ET
FAX Number:	304-558-3970



DataRobot RFP Response:
West Virginia Health and Human Resources

WEST VIRGINIA
Department of
Health & Human
Resources

Sara Marshall
Major Account Executive
sara.marshall@datarobot.com
317-258-8015

DataRobot

CONTENTS

COMPANY INTRODUCTION

Founded in 2012, DataRobot is a privately-held Automated Machine Learning software company (C-Corporation) with approximately 1,000 employees. Headquarters are in Boston, MA (225 Franklin St, 13th Floor), with operations around the world including London, Tokyo, Singapore, Hong Kong and Sydney.

Machine Learning is a type of Artificial Intelligence that allows computers to program themselves based on the data presented, for the purposes of creating predictive models. As an automation of the data science work of creating Machine Learning, DataRobot's platform is included in a technology sector called Automated Machine Learning (AutoML). To date, DataRobot has raised over $1B in capital funding, enabling the company to make strategic investments in R&D and through acquiring companies to complement our platform:

- **May 2017:** DataRobot acquired Nutonian for time series analysis and symbolic regression capabilities valuable in machine learning and genetic programming.
- **July 2018:** DataRobot acquired Nexosis to extend AutoML to the global software development community.
- **February 2019:** DataRobot acquired Cursor, who provides a data collaboration platform which helps organizations find, understand and use data more efficiently.
- **June 2019:** DataRobot acquired ParallelM, who created the machine learning operations (MLOps) category, which helps organizations scale the deployment, management, and governance of machine learning in production using any ML platform on any cloud or on-premise environment.
- **December 2019:** DataRobot acquired Paxata, who helps companies prepare data for AI and other applications.
- **June 2020:** DataRobot acquired Boston Consulting Group's SOURCE AI Technology and entered into a strategic partnership.
- **May 2021:** DataRobot acquired Zepl, a cloud data science and analytics platform.
- **July 2021:** DataRobot acquired Algorithmia, a machine learning operations (MLOps) platform which aims to bring models into production and deliver business value with enterprise-grade security and governance, continuous integration, and accelerated deployment.

DataRobot is a leader in data science, machine learning. Top IT research and consulting companies agree:

DataRobot was named a Visionary in Gartner's 2020 Magic Quadrant for Data Science and Machine Learning Platforms. Gartner writes:

> *"DataRobot leads the charge to incorporate augmented analytics within DSML. It continues to define and demonstrate the use of augmented analytics to engage new types of users in collaboration with traditional roles. DataRobot provides capabilities beneficial to a wide variety of roles including developers, data scientists, statisticians and business analysts."*

Additionally, according to Forrester's 2022 New Wave evaluation, DataRobot leads the pack with a broad set of robust capabilities and is setting the standard of what it means to be an enterprise AutoML solution. In this Forrester report, they write:

> *"Reference customers appreciate the low barrier to entry of adopting the platform and its ease of use ... DataRobot is a solid option for enterprises that want a platform that has tooling for extended AI teams while simultaneously providing collaboration and scale to manage existing use cases and crank out new ones."*

EXECUTIVE SUMMARY

Deploying AI projects at an enterprise scale is currently a big challenge as organizations pursue AI transformation. Automation and augmentation are circuit breakers - together boosting data science productivity and helping organizations to build a core of "citizen data scientists." DataRobot's investments into Automated Machine Learning, Automated Time Series, model interpretability, model deployment, and model management will allowWest Virginia's Department of Health and Human Resources (DHHR) to maximize value by understanding and appropriately applying automation within its processes and workflows.

Through partnering with DHHR's Office of Management Information Services (OMIS), Bureau for Behavioral Health (BBH), and Office of Drug Control Policy (ODCP) to develop and productionize up to 18 AI projects, Team DataRobot will enhance BBH and ODCP's data analysis and modeling capabilities across users and data systems to inform policy and make operational decisions. AutoML will radically boost productivity across information technology staff, contracted data scientists, and epidemiologists to deliver on DHHR's mission and goals.

When augmented by DataRobot, DHHR teams will simplify and maximize the value out of their data and processes. DHHR will be able to incorporate machine learning into their current workflows and build predictive models at scale, allowing for better decision-making across a variety of use cases. With DataRobot, DHHR has the opportunity to democratize data science and include analysts and other staff in their AI-building community. By expanding the pool of people who can use Machine Learning to build and deploy AI systems, DHHR can address a much higher volume of AI opportunities, but also complete AI projects in orders of magnitude faster than traditional approaches.

Coupling AWS GovCloud and DataRobot's AI and Machine Learning as we have done with many other clients, Team DataRobot's proposed solution will enable DHHR to adopt a secure, compliant, and best-of-breed solution that delivers the best possible results while reducing the risk of vendor lock-in. DataRobot can be integrated with numerous other vendors across DHHR's technology stack, including open source tools, PowerBI, Tableau, and SAS to name a few. Team DataRobot's proposed solution includes:

- AWS GovCloud
- Amazon Athena
- AWS Glue
- AWS Lake Formation
- Amazon S3
- Amazon Macie
- DataRobot End-to-End AI Platform (AutoML/AutoTS, MLOps, App Builder)
- Ability to leverage your existing business intelligence (BI) visualization tools

Over the course of the partnership, Team DataRobot will collaborate and enable DHHR to identify, prioritize, and execute on a portfolio of end-to-end AI projects that will generate meaningful impact. At the start of the engagement, DHHR will be assigned an Account Executive, a Strategic Account Manager, a Project Manager, a Customer Facing Data Scientist, and a Field Engineer. Team DataRobot will work with DHHR to understand the specifics of what success means for the duration of the contract. After the vision is documented, Team DataRobot will support DHHR to assess use cases, create Statements of Work, and execute up to 18 AI projects.

As part of Team DataRobot's best practices developed over a decade, Team DataRobot executives will schedule quarterly executive business reviews with DHHR to ensure teams are continually aligned on DHHR vision and goals, provide updates on completed and in flight projects, review outcome metrics and value, and discuss next steps/upcoming projects.

Team DataRobot welcomes and looks forward to having the opportunity to continue partnering with DHHR and contribute to the advancement of its AI evolution.

4.2.2 MANDATORY PROJECT REQUIREMENTS

4.2.2.1 TRAINING

4.2.2.1.1

Vendor shall provide 40 hours of instructor lead, virtual training to up to 10 Agency Staff covering the data cleansing and transformation tools and all of the functions of the Predictive Analytics Software.

The proposed solution will include 40 hours of virtual training for 10 DHHR Agency staff.

Team DataRobot also leverages our enablement arm, DataRobot University for self-paced, on-demand learning and live, virtual instructor-led training. DataRobot University begins with the practical, teaching customers what they need to know to start solving real-world problems immediately. Because most of the technical work is automated by our AI platform, training focuses on other skills customers need to ensure success. DataRobot University premium subscriptions, which allow access to all content (self-paced and virtual instructor-led), are available to DHHR employees at no additional cost through 7/31/2024, as these subscriptions were originally purchased through the agreement with the West Virginia Office of Economic Development and are still accessible. Note that if a contract award extension were to be granted beyond 7/31/2024, DataRobot University subscriptions would be added to the contract at that time.

4.2.2.1.2

Training shall be broken down, at a minimum, into segments for Data Transformation, Predictive Analytics for Developers, and Predictive Analytics for Viewers.

Topics will be sure to include Data Transformation, Predictive Analytics for Developers, and Predictive Analytics for Viewers and cover the full lifecycle of data preparation through model deployment. While all DHHR users will have access to the same capabilities through the DataRobot platform, Team DataRobot recognizes that the term "developers" is synonymous with our term "power users." Team DataRobot will provide training for two separate user groups, understanding that their work on the platform will differ from one another.

4.2.2.1.3

Training shall be interactive and shall be recorded and made available for Agency use as a refresher or to train additional licensees.

These training sessions will be recorded and posted to a private cloud repository. Access will be granted to all identified DHHR users and recognize this may not be limited to 10 unique users over the duration of our partnership.

4.2.2.1.4

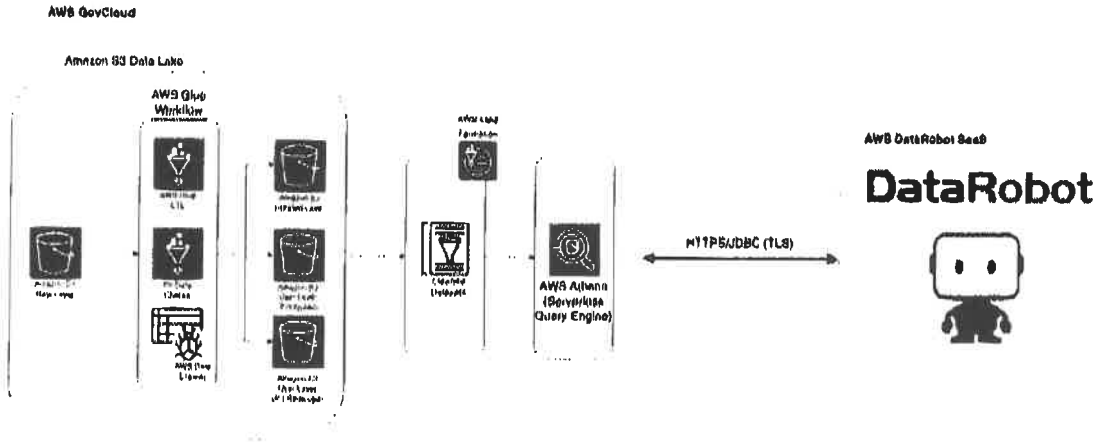
Training shall be provided during regular business hours and will not exceed 8 hours per business day.

These sessions will be coordinated with the DHHR team during business hours and not exceed 8 hours per business day. Sessions will typically range between 1-3 hours, depending on the topic.

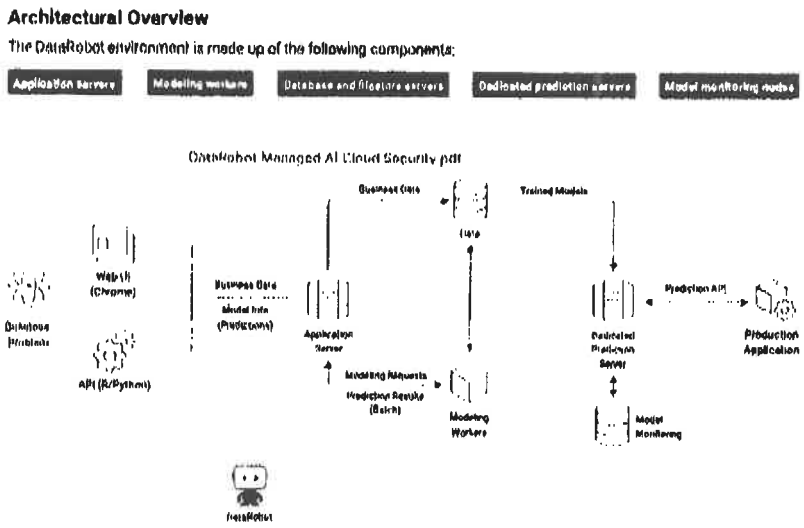
4.2.2.2 Data Repository/Secure File Transfer

Architecture Summary

The Data Repository will sit in AWS GovCloud, be built using the following reference architecture providing up to 2 TB of storage, and be customized to the needs of the DHHR with integrations into DataRobot for data modeling.



The key modules of the DataRobot environment are the application server, data layer, modeling workers, and prediction servers. The following diagram shows a high-level flow of the platform.



The application server houses all of the main administrative components. It handles authentication, project management, and user administration, and provides an endpoint for our API. It also manages the queue of modeling requests made by various projects, which are picked up by the modeling workers – a computing resource that allows DataRobot users to train machine learning models in parallel, and in some cases, also generate predictions. They are also stateless, which allows us to configure them to join and leave the environment on demand.

Trained models are written back into the data layer, and their accuracy is reflected on the model leaderboard through the application server. Trained models can also be deployed to our prediction servers.

Dedicated prediction servers are the most important part of any analytics-based business. They allow real-time decisions to be made quickly and reliably without any concern of failure or delay. Furthermore, key statistics about those predictions and the data provided is returned back to the application server and displayed to users for monitoring the health of the models. The prediction server can also be deployed in an environment disconnected from the DataRobot platform, allowing enterprises to deploy models in segregated networks.

Internal Components
All components within DataRobot are modular in design and can be easily distributed across multiple machines. It is this design that allows DataRobot to scale vertically and horizontally as business demands change. All services are run within Docker containers. This allows multiple instances of certain services to run on multiple machines, providing high availability and resilience in disaster recovery situations.

4.2.2.2.1

Vendor shall host the Agency data in vendor's secure, U.S. based, Cloud repository. The Agency will not provide an environment for the repository.

Team DataRobot will build and host a secure, cloud repository for DHHR in AWS GovCloud.

4.2.2.2.1.1

Vendor will provide adequate Cloud storage and compute resources for 10 Agency users and up to a total of 18 Agency projects, adding resources as necessary to avoid performance degradation.

The cloud storage environment will be built upon a scalable, serverless environment providing adequate compute and storage for DHHR to consume.

The DataRobot Platform is available as a managed cloud service built on top of the Amazon Web Services (AWS) infrastructure. The Proposed Solution includes DataRobot AutoML and AutoTS (supervised learning, time series forecasting and anomaly detection) - 60 workers and 20 model deployments managed by MLOps. With 60 workers, DHHR will have enough computational resources to accommodate 10 users. We generally recommend 6-8 workers per concurrent user; however, we recognize the instances of all 10 users working concurrently will be minimal. Even at 6 workers per user (if all 10 users are working concurrently) the platform will be fully accessible; models will simply be queued until workers in the pool become available. Further, the number of users accessing the platform concurrently for modeling has no effect on the speed of the prediction servers that are used for deployments. The 20 model deployments will satisfy plans of pursuing up to 18 projects over the year. Depending on the project scope, each project is expected to require 1 or more model deployments. Additional workers and model deployments can be purchased throughout the year based on usage and project needs.

4.2.2.2.1.2

Agency staff must be able to securely transfer data and models in formats including, but not limited to those included in 4.2.2.4.2 (.csv, .tsv, .dsv, .xls, .xlsx, .sas7bdat, .geojson, .gz, .bz2, .tar, .tgz, .zip), to the Agency SFTP for use with other Agency software.

The cloud data repository will support structured, semi-structured, and unstructured data including but not limited to the following data types: .csv, .dsv, .txt, .tsv, .xls, .xlsx, .sas7bdat, .geojson, .bz2, .gz, .zip, .tar, .tgz.

Data sent from the Data Repository will be securely transmitted over HTTPS to the DataRobot platform.

The DataRobot Managed Cloud stores all relevant data within a private AWS S3 bucket. This bucket is encrypted for maximum security.

4.2.2.2.2

Data repository shall include industry standard antivirus and antimalware protection. Vendor must name the products utilized in their response.

Team DataRobot will use Antivirus for Amazon S3 from the AWS Marketplace.

DataRobot implements risk-based and standards-based security protocols to secure both our services and customer data. As a part of our comprehensive security program, our managed cloud service is SOC 2 Type II and ISO 27001 certified by an independent third-party auditor to ensure compliance with industry standards and best-practices for information security, corporate controls, and software development. Endpoint security controls protect DataRobot endpoints from a variety of threats including malware, vulnerabilities, and other threats to data stored on endpoints in the event of compromise or theft. Endpoints include user laptops, workstations, and corporate servers.

Please refer to the attached **Appendix A**: DataRobot Trust Package for more information regarding our security processes, policies, and positions.

4.2.2.2.3

Data in the repository shall be encrypted both at rest and in transit.

The cloud data repository will support encryption at rest and in transit.

Data is encrypted at-rest and in-transit (see page 2 of "DataRobot Security Controls Report" in Trust package). The Managed AI Cloud service uses the AWS S3 file system, which is secured with Server-Side Encryption with Amazon S3-Managed Keys (SSE-S3). On top of file system encryption, all data transferred to and from AWS S3 is encrypted in transit using TLS 1.2 for metadata stores, along with password enablement. When using the DataRobot Managed Cloud environment, data sent to be trained or scored is transmitted encrypted with TLS 1.2 over the public internet.

4.2.2.2.3.1

All mechanisms used to encrypt data shall be FIPS 140-2 compliant and operate using the FIPS 140-2 compliant module (Standards available on the National Institutes for Standards and Technology (NIST) Website – https://csrc.nist.gov/publications/detail/fips/140/2/final). Vendor must name any products utilized to provide encryption.

DataRobot uses AWS Key Management Service (KMS) to manage encryption keys. KMS is FIPS 140-2 compliant as of 2018.

(https://aws.amazon.com/blogs/security/aws-key-management-service-now-offers-fips-140-2-validated-cryptographic-modules-enabling-easier-adoption-of-the-service-for-regulated-workloads/)

4.2.2.2.3.2

Storage devices where data has resided must be securely sanitized according to MARS-E MP-6 Media Sanitization security prior to use. A guidance document is available at the Centers for Medicare and Medicaid Services website (https://www.hhs.gov/guidance/document/minimum-acceptable-risk-standards-exchanges-mars-e-20).

"In developing MARS-E v. 2.0, CMS relied on the CMS Acceptable Risk Safeguards (ARS) v. 2.0, as the basis for the security and privacy control requirements. The CMS ARS is based on National Institute of Standards and Technology (NIST) Special Publication (SP) 800-53 Rev 4, Security and Privacy Controls for Federal Information Systems and Organizations (CMS, MARS-E Document Suite 2021)."

The Cloud Data Repository will reside within AWS GovCloud. More information regarding the compliance requirements satisfied by AWS GovCloud can be viewed on the GovCloud website at https://aws.amazon.com/compliance/programs/. AWS Cloud infrastructure and services have been validated by third-party testing performed against the NIST 800-53 Revision 4 controls, as well as additional FedRAMP requirements.

As it pertains to DHHR data: PII, PHI, PCI and other sensitive or confidential data will be rejected back to DHHR from the data repository before information is permitted on DataRobot's Managed AI Cloud. DataRobot does not need and will not want access to any regulated personal data and will therefore not be considered as a Data Processor under any relevant data protection legislation.

4.2.2.2.4

Data repository shall include a perimeter firewall. Vendor must identify the firewall that is used.

The appropriate firewall and security requirements will be configured using native AWS services for the cloud data repository.

DataRobot implements its WAF (Web Application Firewall) to protect the DataRobot application from malicious web traffic. The local firewall is also enabled and required on servers that store sensitive information. DataRobot also uses traffic filtering as a network security control.

4.2.2.2.5

Data will be stored in at least two geo redundant locations making it improbable that a single event, whether naturally occurring or manmade, will impact both locations. In the event operations are interrupted at the primary data center, Agency operations will be shifted to the secondary location within 4 hours.

Team DataRobot will leverage several AWS Availability zones across several regions to ensure redundancy and that services are not impacted due to unexpected events. Please refer to the attached **Appendix A** DataRobot Trust Package for more information regarding our security processes, policies, and positions.

4.2.2.2.6

Vendor shall scan incoming data for fields that appear to contain Personally Identifiable Information (PII) or other sensitive data types and reject flagged files back to the Agency to verify no sensitive data is included.

Team DataRobot will set up an automated process within AWS to scan for PII data.

4.2.2.2.7

Vendor shall acknowledge that all data in the repository is the property of the Agency and will be provided to the agency upon request. Data in the repository at the end of the contract period will be provided to the Agency in a mutually agreeable format and upon written notice by the Agency, all copies in the possession of the vendor will be destroyed with a certificate of data destruction provided to the agency.

DataRobot provides two options for deleting projects. Owners can delete projects so that they are removed from active project management listings. Completely removing the project, including the data used to build it, requires an administrator. As Managed AI Cloud users, DHHR users wanting to permanently remove data would ask their primary DataRobot contact to file a ticket with DataRobot Support and request the data be removed. Confirmation of deletion would occur in response to the Support ticket.

4.2.2.2.8

Agency will upload cleansed and transformed data to the Agency SFTP Server and notify the vendor by email when it is available. Vendor will move the data from the SFTP server to the data repository.

Team DataRobot will build an automated data pipeline that will upload data to the SFTP server after being cleansed and an email notification will be generated to DHHR.

4.2.2.2.9

Vendor will certify that the hosted cloud environment satisfies MARS-E privacy controls (available at https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/3-MARS-E-v2-0-Catalog-of-Security-and-Privacy-Controls-11102015.pdf), including privacy training and awareness, and rules of behavior.

The cloud data repository will be hosted in AWS GovCloud. More information about the regulatory compliances that GovCloud satisfies can be found on the GovCloud website: https://aws.amazon.com/compliance/programs/. AWS Cloud infrastructure and services have been validated by third-party testing performed against the NIST 800-53 Revision 4 controls, as well as additional FedRAMP requirements.

DataRobot complies with international data privacy regulations, including GDPR and CCPA. See our Privacy and Personal Data policy here: https://www.datarobot.com/trustcenter/privacy-and-personal-data/ Note that the DataRobot Managed Cloud environment does NOT support PII ingestion and is not certified to maintain and protect PHI (Personal Health Data) or PII (Personally Identifiable Information).

The DataRobot Privacy Policy governs how DataRobot collects and uses personal information gathered online through the DataRobot website and is available at this URL: https://www.datarobot.com/privacy/.

4.2.2.2.10

Vendor agrees that the hosted cloud environment will be available to Agency staff for data transfer and data modeling 99% of the time, 24 hours per day, 7 days per week, with the exception of scheduled downtime.

Team DataRobot agrees that the hosted cloud environment will be available to Agency staff for data transfer and data modeling 99% of the time, 24 hours per day, 7 days per week, with the exception of scheduled downtime.

DataRobot's full support policy can be found
https://www.datarobot.com/wp-content/uploads/2022/02/DataRobot-Support-Policy-Live-14-Feb-2022.pdf.

4.2.2.2.10.1

Vendor proposal shall include maintenance windows and scheduled downtime which shall occur in off-peak hours, between 8:00 p.m. and 6:00 a.m., Eastern Time (ET) Monday through Friday or on Saturday, Sunday or State Holidays.

Please refer to the full support policy referenced above in subsection 4.2.2.2.10

4.2.2.2.10.2

Vendor shall provide a system downtime report delineating both scheduled and unscheduled downtime for the month, with each monthly invoice for cloud services.

Scheduled maintenance and any downtime in the last 90 days can be viewed at https://status.datarobot.com/. All maintenance is included in the software license fee. DataRobot fully warrants its product. In addition, DataRobot offers 24x7 Sev1 break/fix support. Resources will be available globally 24x7 until a resolution or workaround is in place. All other Severity levels are supported Monday through Friday during normal Business Hours. DataRobot's support offices are located strategically around the world, allowing us to provide "follow the sun" support. Hours of operation for each location are based on local standards.

DataRobot will use commercially reasonable efforts, commensurate with the severity of the error, to correct any malfunction, defect or non-conformity ("Error") in the operation of the Software so that it will substantially perform in accordance with the product documentation. The customer shall conduct reasonable and adequate research with respect to any claimed Error or related issue prior to contacting DataRobot for assistance. Customer and DataRobot will work together to resolve requests for information, documentation, technical assistance and other assistance regarding any such Error. Each reported Error will be logged and tracked by DataRobot, and will remain open until the issue is resolved. The customer may designate the priority level when submitting an Error, however, DataRobot reserves the right to adjust priority in its reasonable discretion. Any support for Upgrade(s) will be designated as a Severity 3. DataRobot shall give DHHR no less than 14 days' notice of any scheduled maintenance and as much notice as possible for any other maintenance. Notice of scheduled maintenance will be given at https://status.datarobot.com/. DHHR can subscribe to email updates to the page using the subscribe function on the page. While DataRobot does not contract to specific SLAs, we do endeavor to maintain at least 99.9% availability in our Managed AI Cloud meeting DHHR's requirement of 99% availability.

SEVERITY	INITIAL RESPONSE TARGET	UPDATE FREQUENCY TARGET
Severity 1	Within 1 hour	Continuous effort with written updates every 4 hours
Severity 2	Within 2 Business Hours	Updated every Business Day
Severity 3	Within 8 Business Hours	Updated every 3 Business Days
Severity 4	Within 2 Business Days	N/A, feature request

4.2.2.2.10.3

Vendor agrees that for any month unscheduled downtime is greater than 1% but less than 2.51%, Agency may deduct 2.5% from the total due on the monthly invoice. If unscheduled downtime is equal to, or greater than 2.51%, Agency may deduct 5% from the total due.

Please refer to DataRobot's Service Availability Policy:
https://www.datarobot.com/wp-content/uploads/2022/02/DataRobot-Availability-Policy-LIVE-14-Feb-2022.pdf
. Section 2, Remedies for Missing Quarterly Uptime, is captured below. The link above provides full policy.

2. **REMEDIES FOR MISSING QUARTERLY UPTIME**

2.1 If Quarterly Uptime falls below 99.95% in a calendar quarter, DataRobot shall pay Customer a service credit as follows ("**Service Credit**"):

Availability	Service Credit
97.0% - 99.94%	5 percent of the fees for the affected Solution for the applicable calendar quarter
95.0% - 96.9%	10 percent of the fees for the affected Solution for the applicable calendar quarter
Less than 95%	20 percent of the fees for the affected Solution for the applicable calendar quarter

4.2.2.3 DATA MODELING PROJECTS

4.2.2.3.1

Agency will initiate project requests by preparing a data set and uploading to the SFTP server. Vendor shall move the dataset to the hosted cloud repository. Agency staff will perform preliminary data modeling in the cloud before initiating a project with the vendor.

Team DataRobot will move the dataset to the hosted cloud repository from the SFTP server and scan data for PII per 4.2.2.2.6 requirement. We acknowledge Agency staff will perform preliminary data modeling in the cloud before initiating a project with the vendor.

4.2.2.3.2

Within 2 business days of a request to initiate a project, Vendor shall schedule a meeting with Agency staff to occur within 5 business days. Agency and vendor will determine the project scope including desired outcomes, number of models desired, and a not to exceed estimate of project duration

(expressed in hours of support required per week). Within 2 business days after the meeting, the vendor will provide a draft project scope for Agency approval.

Team DataRobot acknowledges that DHHR will submit a request to initiate a project. Team DataRobot will respond within 2 business days of request and schedule a meeting within 5 business days. Project scoping will contain desired outcomes, number of models desired, and project duration. Turnaround time of scope drafts will be 2 business days after the meeting.

4.2.2.3.3

Upon receipt from the Agency of an approved project scope vendor shall begin providing up to 5 hours of data scientist support per week to the Agency project staff at the data scientist billable rate proposed in the RFP response until the scope of work is satisfactorily completed. Vendor or Agency may request fewer support hours per week, spreading the total hours over a longer period of time but any such modification shall require mutual agreement of the parties in the project scope.

Team DataRobot will provide up to 5 hours of data science support per week per project (up to 6 concurrent projects) to DHHR project staff at the billable rate proposed in RFP until scope is satisfactorily completed. Modification requires mutual agreement of the parties.

4.2.2.3.4

Vendor agrees that the project duration in the scope of work is a not to exceed estimate and the hours billed shall represent actual hours worked.

Team DataRobot agrees that the project duration will not exceed estimate and the hours billed shall represent actual hours worked.

4.2.2.3.5

Agency may request changes to the scope of work resulting in a modified scope of work. Vendor shall prepare a new estimate of required support for the modified Statement of Work. Changes in scope that add no more than 25% to the project duration shall be considered a project change and added to the maximum billable hours for the project. A scope change that adds greater than 25% shall be considered a new project.

Team DataRobot will prepare new estimates of required support for any modified Statement of Work due to changes in scope or issues with dependencies/assumptions. Changes that require up to 25% more billable hours will be considered a project change. If more than 25% billable hours are required for the Statement of Work, Team DataRobot recognizes that this scenario will be considered a new project.

4.2.2.3.6

Vendor shall support up to 18 total projects during the 12 month life of the contract including up to 6 projects concurrently.

Team DataRobot will support up to 18 total projects during the 12 month life of the contract including up to 6 projects concurrently.

4.2.2.3.7

Agency may adjust the priority of projects, placing a lower priority project on hold to keep the number of concurrent projects to six or fewer. Vendor shall accommodate the Agency priorities.

Team DataRobot will accommodate the prioritization of projects based on DHHR direction and feedback. As part of Team DataRobot's best practices developed over a decade, Team DataRobot Strategic Account Manager and Project Manager will schedule quarterly executive business reviews with DHHR to ensure teams are continually aligned on DHHR vision and goals, provide updates on completed and in flight projects, review outcome metrics and value, and discuss next steps/upcoming projects.

4.2.2.3.8

Upon contract award, the Vendor shall designate one primary contact and at least one backup that will be the initial point of contact for all project engagements under this contract. Only projects properly initiated with the Vendor point of contact are valid projects under the contract.

See Appendix B for resumes. The Strategic Account Manager (SAM) will be Vendor's primary point of contact, backed up by the Project Manager (PM). Only projects initiated through the SAM or PM will be considered valid under the contract.

4.2.2.4 REQUIRED SOFTWARE

4.2.2.4.1

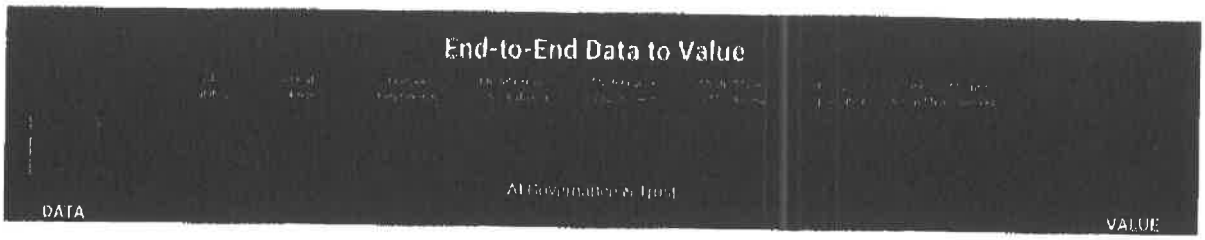
Vendor shall list in their technical proposal ALL of the software the State Agency will need in order to satisfy the stated goals and objectives and to meet all of the mandatory requirements within the RFP.

Team DataRobot outlines the proposed solution architecture under Subsection 4.2.2.2 Data Repository/Secure File Transfer, Architecture Summary. If awarded to Team DataRobot, DHHR will be procuring the AI Platform and paying a data repository hosting fee. The repository hosting fee covers the cost for a managed AWS environment which includes AWS services and the engineering hours to automate, manage, and maintain the data pipelines.

In Summary, DHHR will need to only order the DataRobot AI Platform.

The DataRobot AI Platform is a consolidated, unified platform which offers cutting-edge capabilities from the point of ingesting data from the secure, compliant data repository described in this response, through to empowering decision management with AI-Powered Applications. Team DataRobot is geared towards an automation-first approach and has been used in a variety of organizations across the globe.

The DataRobot AI Platform will accelerate DHHR's AI success by augmenting the technical team and user community already in place with cutting-edge machine learning technology. By incorporating the knowledge, experience, and best practices of the world's leading data scientists, DataRobot delivers unmatched levels of automation, accuracy, transparency, and collaboration to help DHHR transform into an AI-driven enterprise. Enterprise model lifecycle features of DataRobot highly valued by our Tier 1 customers include: separate engines for model training and model scoring; simple, one-click model deployment; model management; model monitoring in production; automated retraining.



4.2.2.4.2

Predictive analytics software proposed in response to this RFP must have the capability to process text data via natural language processing and must handle multiple file formats including, but not limited to (.csv, .tsv, .dsv, .xls, .xlsx, .sas7bdat, .geojson, .gz, .bz2, .tar, .tgz, .zip). The software must be able to export data in formats that are compatible with popular data visualization software including, but not limited to Tableaux and Microsoft Power BI.

Below we have provided details on the primary components supporting the end-to-end machine learning lifecycle. Please see specifically the section titled 'DataRobot's AI Catalog' confirming that the system can support this broad range of formats'. Please also see section 'Integrations' confirming that we are compatible with popular data visualization software

DataRobot's AI Catalog

DataRobot facilitates user groups and teams to collaborate from a centralized repository of data assets called the AI Catalog. End users are able to upload structured and unstructured data into the AI Catalog from a variety of data sources (SQL-type databases, URLs, local files, zipped directories of images) through either a graphical user interface or via the DataRobot API.

The AI Catalog supports a range of data formats from structured datasets (e.g. .csv, .dsv, .txt, .tsv, .xls, .xlsx, .sas7bdat, .geojson, .bz2, .gz, .zip, .tar, .tgz) through to unstructured data (image data). Free text data can be automatically handled as a feature in a structured dataset, which means descriptions and notes can be automatically leveraged for modeling without text mining or other pre-processing. To call out image data in more detail, powerful use cases can now be delivered using Visual AI (https://www.datarobot.com/platform/visual-ai/), a key DataRobot feature. As images can be incorporated as standard predictive features without special manual pre-processing, they can be combined with typical structured data - this combination of augmenting image data with structured data is truly a differentiator in the AI technology landscape. We have also introduced the ability to exploit location-based data in the modeling datasets using latitude/longitude pairs or geojson files. The main advantage is that users can gain insight and improved prediction accuracy by incorporating spatial data (e.g., nearest neighbors) into the predictive models they build and use.

Exploratory Data Analysis, Data Quality, Feature Engineering, and Feature Discovery

It is fully understood that a large proportion of any analytics project can be spent in the data preparation phase. A common statistic places that proportion at roughly 80%. In response to this challenge, DataRobot has made a large investment in reducing the manual effort needed in this part of the modeling lifecycle.

When a DHHR user starts a project from a dataset in the AI Catalog, the DataRobot platform provides automation and guidance to ensure that data is handled appropriately and supports the users in their modeling efforts. The first step is automatically doing an Exploratory Data Analysis, which helps users uncover trends in the data through histograms, feature typing, and other descriptive statistics.

Second is data quality. DataRobot automatically detects low information, duplicate, or target leakage features which might undermine model accuracy. DataRobot automatically assesses Data Quality for ingested data, including for inliers, outliers, disguised missing values, excess zeros, target leakage, and missing and duplicate images (for models using Visual AI). Both of these steps remove noise from the data or prompt users to potential issues in their modeling data.

Next, DataRobot supports an automation-first approach to feature engineering. DataRobot creates smart features by default, such as day of month and day of week for date features. This is a general best practice in modeling with date features. For time series models, DataRobot automatically creates lagged features based on user specifications, saving days or weeks of manual creation and testing of moving average features generally used in forecasting. DataRobot also creates interaction features within a dataset using the "Search for Interactions" functionality, which looks for features that explicitly mediate or moderate one another. This allows content experts to use their expertise - DataRobot's automation functionality is there to accelerate the process of feature engineering and building models.

Finally, DataRobot reduces the manual effort of feature engineering through Automated Feature Discovery. With this feature, users can automatically discover and generate new features from multiple datasets and no longer need to perform the feature engineering manually to consolidate multiple datasets into one. From the primary dataset, DataRobot guides the user through creating relationships to additional datasets, called secondary datasets. The end result is a multitude of additional features that will be used to train more accurate models and generate predictions.

These features are automation-first, but they are not meant to replace content experts. DHHR SMEs have the ability to create custom feature lists beyond the automatically-generated Informative Features list and to apply mathematical transformations to features such as logs, square roots, etc. Any feature engineering done in DataRobot during the modeling phase is automatically applied to incoming rows during the scoring phase of the project. Transformations are fully documented and Lineage is provided.

Model Creation and Validation

DataRobot is the category-creator and leader in the Automated Machine Learning space. Two core tenets of DataRobot are to democratize data science while improving existing data science capabilities. Consequently, after data preparation, DataRobot allows less experienced users to immediately begin the automated machine learning process with the push of a button (guardrails are in place behind the scenes). More experienced users, e.g. data scientists, have a wide range of flexibility, control, and configuration options available to customize the modeling setup.

DataRobot can solve a wide array of machine learning problems, in particular:

1. **Binary Classification problems**: The most common type of supervised learning, the goal of binary classification is to estimate the relative probability of two different and exclusive outcomes (generally 0/1 or "yes"/"no") for some new input. For instance, "Will patient B be readmitted to the hospital within 60 days?"
2. **Multi-class classification problems**: An extension of binary classification is multi-class classification, which DataRobot can currently handle for problems up to 100 classes. In this type of model, the outcome is the probability distribution over more than two possible options. For instance, "Which of 20 common diseases is likely afflicting a tree, given an image of one of its leaves?"
3. **Regression problems**: Regression problems are supervised learning problems in which the target feature is continuous or numeric rather than a probability. For instance, the sales price of a home or the expected duration (in days) that a construction project will take to complete.

4. **Time Series:** Time series forecasting is a specific type of modeling that relies on derived time-based features (e.g., moving averages or lags) within a user-specified period of time to predict future values of the outcome of interest. The most common application is demand forecasting - given daily demand data for each product in a store for the past 2 years, predict demand each day for the next 30 days. DataRobot provides an advanced set of algorithms suited specifically for time series predictions. They range from simplistic approaches like the ARIMA method to complex ones like the Eureqa and XGBoost.
5. **Anomaly detection:** Anomaly detection is a type of unsupervised model, meaning there is no specific target to try and predict. Instead, anomaly detection algorithms identify abnormal patterns (outliers) non-conformant to expected behavior.

After the modeling process is initiated, DataRobot will automatically proceed through the following steps:

- Perform pairwise correlations between features and univariate correlations between each feature and the target feature.
    - Feature Association Matrix: Users can explore - in an intuitive visual interface - correlated features, which can help guide them in refining their dataset to improve model accuracy and/or performance.
- Analyze the distribution of the target feature and the predictive features to determine the most appropriate suite of "blueprints" to run, typically anywhere between 30-40 blueprints.
    - A blueprint encapsulates the collected expertise of hundreds of DataRobot's data scientists (some of whom have been ranked #1 in the world on Kaggle, a global data science competition)
    - Each blueprint contains both data preprocessing steps, i.e., missing value imputation, and one to many machine learning algorithms based on the type of data available
        - E.g. If there is numeric, text, and categorical data, a blueprint might include a step to pre-process the free text feature into a word vector, impute any missing values in the numeric features, and encode the categorical features in numeric form. Then the blueprint would feed that transformed data into a sequence of machine learning algorithms to generate a final result. All of this is done automatically and with no manual intervention by the user.
    - The library, or "Repository", within DataRobot contains hundreds of blueprints that can be chosen based on the problem type.
    - DataRobot employs a unique strategy called "survival of the fittest" to iteratively test and tune the 30-40 blueprints against increasing subsets of data. This helps DataRobot to rapidly build the most accurate models on large datasets.
    - Advanced users are able to run additional models from the Repository and can also manually tune any hyper-parameter. They can also adjust or tweak DataRobot blueprints with Custom Blueprint functionality.
        - Every step in a DataRobot blueprint comes with extensive documentation, including links to the original academic white papers when available.
- Another core tenet of DataRobot is model transparency. DataRobot provides a number of unique features to enable citizen data scientists to understand a model:
    - Feature Importance displays the most important features for each model, e.g. showing that reliable transportation and household income are the most important predictors for no show appointments.
    - Feature Effects allows users to investigate the impact of altering a particular feature on the model's prediction.
    - Prediction Explanations provide a list of row-level explanations, which provide a specific rationale for every prediction.
    - Other ways to interpret a model include linear coefficients, word clouds for text features, high level rules, etc.

- DataRobot will automatically fit and rank models according to a suggested Accuracy metric. Users can then deploy any of the models to MLOps with only a few clicks.
- DataRobot provides robust Python/R APIs for users to develop or import custom models as needed.

Predictions and Machine Learning Operations (MLOps)

Prediction Methods

DataRobot offers several methods for getting predictions on new data. DataRobot provides the ability for generating predictions through a range of different methods. These can be used interchangeably depending on the business use case, and include:

- Real-time predictions: Make real-time predictions by connecting to HTTP and requesting predictions for a model via a synchronous call. Predictions are made after DataRobot receives the request and immediately returns a response.
- Batch predictions: After deploying a model, you can make batch predictions via the UI by accessing the deployment, or use the Batch Prediction API. These batch predictions can be scheduled in advance and configured to write predictions back to a database connection, such as to Tableau, Snowflake, or SQL Server.
- Portable predictions allow you to execute prediction jobs outside of the DataRobot application. There are three main kinds of portable predictions:
    - Scoring Code: You can export Scoring Code from DataRobot in Java or Python to make predictions. Scoring Code is portable and executable in any computing environment. This method is useful for low-latency applications that cannot fully support REST API performance or lack network access.
    - Portable Prediction Server: The Portable Prediction Server (PPS) is a remote DataRobot execution environment for DataRobot model packages (MLPKG files) distributed as a self-contained Docker image. It can host one or more production models. The models are accessible through DataRobot's Prediction API for predictions and Prediction Explanations.
    - DataRobot Prime: DataRobot Prime generates fast Python or Java Scoring Code that can be run anywhere with no dependencies. Once created, you can export these models as a Python module or a Java class, and run the exported script.

Once a model is deployed to a prediction server managed by DataRobot, users can make predictions via the API and monitor and manage the deployment with our full suite of monitoring capabilities.

Model Monitoring

DataRobot MLOps offers a "single pane of glass" for production models, simplifying and streamlining the maintenance of deployed models. MLOps offers user an interface that tracks and displays visually:

- Data Drift: the changes in production data distributions for influential factors in order to recommend whether the model needs to be refreshed with more recent data;
- Model Accuracy: the decrease of performance and accuracy of models over time,
- Service Health: the latency, volume and health of prediction requests; and
- Bias and Fairness: the relative performance of your production model against any relevant fairness metrics or protected classes defined for the use case.

This tracking is true regardless of where the model is being executed - whether it was built and deployed on DataRobot, built in an open-source tool like Python and deployed on DataRobot, or built and deployed in an

entirely separate environment using DataRobot MLOps agent functionality. MLOps allows users to not only view model performance across the entire organization through a single, centralized dashboard, but also provides centralized model lifecycle management capabilities.

Custom Inference Models

DataRobot's MLOps solution can serve as a centralized repository for all your models across DHHR, regardless of whether they were originally trained in DataRobot or are custom inference models trained externally using other frameworks (e.g. R and Python). By uploading a model artifact, you can create, test, and deploy custom inference models to a centralized deployment hub. DataRobot supports models built with a variety of coding languages, including Python, R, Scala, and Java.

Champion / Challenger Structure

MLOps allows users to build their own challenger models or use our industry-leading automated machine learning product to build and test them for you. Track how the challenger models are performing against the model in production (the champion) and easily swap the champion model if the challenger proves to be more accurate. This process of constant evaluation enables you to avoid surprise performance regressions in a dynamic and highly volatile environment.

Integrations

Through DataRobot's flexible REST API routes, DataRobot has the ability to connect and surface insight into many other enterprise tools. These include BI tools like Tableau, PowerBI and Qlik, and RPA tools like Automation Anywhere, UiPath and Blue Prism. Furthermore, DataRobot can support direct prediction write-back with Snowflake, Microsoft SQL Server and Tableau. With these seamless integrations, DataRobot can easily complement your existing technology ecosystem.

AI-Powered Applications

Another challenge organizations struggle with is with orchestrated data-driven decision support.

To address this challenge, DataRobot has developed a suite of decision-support tools called AI-Powered Applications. These are incredibly flexible, but include starter templates for an Optimizer app, a Predictor app and a What-If app. These applications allow front-end business users to leverage DataRobot's predictions directly in decision making without having a DataRobot license or building models themselves. The DataRobot users creating these apps do not need a background in coding or app hosting, making this a scalable solution for incorporating data into business decisions.

Governance and Approval Workflow

DataRobot MLOps establishes a framework in which you can maintain discipline and control over your AI projects.

- The Humble AI feature inspires trust in every single prediction, by allowing users to set custom override rules or to flag decision-makers when it encounters particularly difficult or unique cases. MLOps also enables you to comply with government regulations and reduces your risk by providing secure and governed access to all your production models with tightly controlled approval workflows for the deployment process and the implementation of any proposed changes.
- Model Approval Workflows - Maintain thorough reviews of model updates with less tedious manual work using customizable review cycles and approval workflows. MLOps ensures only those who are authorized can update and publish new models while keeping everyone else in the loop.

- ML Audit Trail and Logging - For regulatory compliance purposes, MLOps preserves a full history of prediction activity and any model updates so that you always know what model was created, used, and updated, when by whom.

Solution Differentiators

As noted in the analyst reports, DataRobot is the leader in automated machine learning and enables citizen data scientists to build models easily. Our key differentiators are:

1. A customer-first engagement model with dedicated data science, field engineering, account management, and customer success resources. Our key success metric is operationalizing machine learning models into end-user workflows in order to generate significant business value. Team DataRobot will provide both consultative guidance and technical support at every stage of the end-to-end AI journey. From initial use case ideation to model development, deployment, and the ongoing maintenance and management of your deployed models - we have dedicated professionals that we pair with world class software to help you achieve your ultimate vision of success.
2. We democratize data science by automating best practices including feature engineering, model building, model evaluation, model interpretability, model deployment, and ongoing model management. We will provide robust training for our product through DHHR's 40 hour customized training and our DataRobot University offering. This ensures that DHHR team members rapidly gain the ability to execute data science projects independently.
3. Robust data integration and ingestion capabilities that enable DHHR to leverage multiple data types (natural language, tabular, image, geospatial, etc.) to build robust models. Our REST API also gives DataRobot users the ability to feed predictions and the explanations behind those predictions to a wide array of end-user solutions and business intelligence tools/dashboards for consumption (including but not limited to Microsoft PowerBI, Tableau, Qlik, EPIC, etc.).
4. We are the only automated machine learning vendor that provides robust model management support to ensure a continued ROI. Our acquisition of ParallelM, the leading machine learning ops company, only enhances our existing capabilities.

4.2.2.4.3

Vendor will clearly identify any software that is proprietary and will explain the basis for software licenses including whether the licenses are named user licenses or concurrent user licenses; whether licenses are annual or perpetual; any requirements requiring software to be under vendor support contracts; etc.

The proposed solution includes the purchase of a platform instance AutoML/AutoTS consisting of 60 workers and 20 model deployments on the MLOps platform. This license is annual and includes unlimited user seats across DHHR. Please see www.datarobot.com/legal/MSA for more detailed information on our standard terms.

4.2.2.4.3.1

Vendor proposal shall indicate whether licenses are transferable (from an Agency staff member leaving the project to a new staff member) and whether and, how a license might be upgraded during the license term, for instance from a view only license to a license with full access to SW features.

DHHR will have the ability to provide access to 10 users outlined in the RFP and any additional personnel within DHHR that it deems necessary. Licenses do not need to be upgraded during the license term. All users granted access by DHHR will have access to the full software features. Section 4.2.2.2.1.1 describes the sizing of the platform based on users. DHHR will have enough computational resources to accommodate 10 users. Licenses

are transferable between DHHR employees, or additional additional capacity (more workers) can be purchased if more users are added. Similarly, more model deployments can be purchased if scoped projects require additional models above and beyond 20 deployments.

4.2.2.4.4

Vendor will clearly identify any required third party software, if Vendor is an authorized distributor of such third party software or if the Agency will have to procure their own licenses. (NOTE: Where an existing Agency or Statewide Contract includes the required third party software, Agency reserves the right to purchase from the existing contract rather than from the Vendor.)

Team DataRobot's proposed solution only requires the purchase of the DataRobot platform.

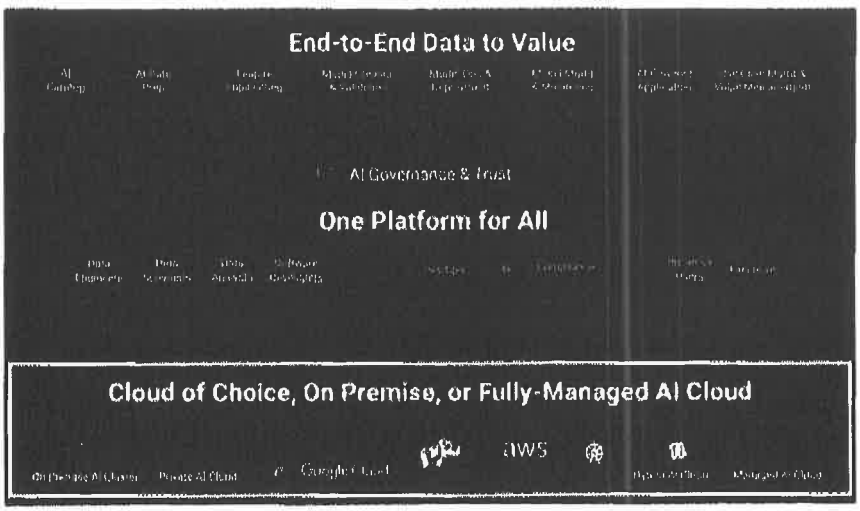
4.2.2.4.5

Vendor will address their approach to SW version and release updates (including bug fixes). The response should include details regarding what updates are required vs. optional; the amount of notice the Agency will be provided for routine updates; the amount of notice the agency will be provided for bug fixes; etc.

Current DataRobot version is 8.0. DataRobot releases a new version every week within the cloud. While our product roadmap is not available for distribution, DataRobot has been and is committed to bringing the most advanced, cutting-edge machine learning capabilities to the market.

DataRobot Managed Cloud is operated in a Blue/Green (active/passive) deployment mode, which ensures no client impact during upgrade. The cloud environment is updated weekly and involves failover between the two sites. Disaster recovery from backups is tested weekly for metadata. Client data is held on a secure, encrypted S3 bucket in AWS which provides eleven nines durability.

Additional details concerning notice of maintenance (including addressing bug fixes) are covered in subsections 4.2.2.2.10 / 4.2.2.2.10.1 / 4.2.2.2.10.2.



4.3. QUALIFICATIONS AND EXPERIENCE

4.3.1 Qualifications and Experience Information

4.3.1.1

Vendor should demonstrate that they have provided predictive analytics SW, including machine learning and artificial intelligence to companies and agencies in the United States for a minimum of five years prior to this bid opening.

DataRobot was founded in 2012. The SaaS market offering has been available for over 6 years. In the 6 years it's been available, over 1000 of the world's most recognised brands like Black and Decker, Toyota, Standard Chartered Bank, Monsanto Bayer, U.S. Government - federal and state entities and so many more organizations trust DataRobot every day to deliver real business value from AI. DataRobot customers include a third of the Fortune 50, with companies across all industries, as well as highly regulated industries, including: 40% of Top Healthcare Companies and 70% of Top U.S. Banks.

Since DataRobot's launch, over 1 trillion predictions have been made, 2.5 billion models have been built, and 1 million projects have been supported by a team of dedicated customer facing data scientists.

DataRobot is the category creator of automated machine learning. Below are recent 3rd party accolades of DataRobot:

- DataRobot recognized as a Leader in the Forrester Wave: AI/ML Platforms, Q3 2022
- DataRobot recognized as a Representative Vendor in the May 2022 Gartner® Market Guide for Multipersona Data Science and Machine Learning (DSML) Platforms report.
- DataRobot recognized as a Category Leader by AI Forum in the AI Quadrant 2022 for Enterprise Machine Learning
- DataRobot recognized as a Leader in the Everest Group MLOPs PEAK Matrix Assessment 2022 and received the highest vision and capability score
- DataRobot is #26 on the 2022 Cloud 100 by Forbes

4.3.1.1.1

Vendor should provide the current release of the SW being proposed and comment on the maturity level of that release.

The current version of DataRobot Managed Cloud is version 8.0. DataRobot pushes weekly updates to our cloud environment, as well as a more significant quarterly release of new functionality. Each release is thoroughly vetted by a formal change management process, and robust validation, testing, and enterprise-grade quality assurance programs. For example, all code commits are subjected to more than 20,000 unique tests to compare known inputs to saved expected outputs. Results are analyzed by data scientists, data engineers, and software developers with significant data science expertise. If any change causes a deviation of more than six decimal points, the code will not be accepted.

In addition to the rigorous testing for each release, DataRobot blueprints (the logic that drives the thousands of tasks that go into a machine learning project) are automatically tested every night using a series of datasets that are carefully curated to utilize every major feature, and simulate the wide variety of use cases and dataset types

that DataRobot would routinely encounter in a production environment. Any regressions in key metrics are carefully reviewed by our data science team.

4.3.1.2

Vendor should provide a staffing plan that will clearly support the requirements enumerated in the RFP. The plan should include, at a minimum, resumes for proposed staff along with copies of certifications or degrees applicable to this contract; descriptions of past projects completed; project manager name and contact information for past projects; and, customer name and contact information for each project cited.

Please refer to **Appendix B** (also cited in section 4.2.2.3.8) for Strategic Account Manager and Project Manager profiles. **Appendix C** includes resumes containing profiles for Data Scientist and AI Architecture/Engineer. Due to Personally Identifiable Information that is not publicly available, contact information and further project details can be provided upon request. Customer point of contact information for government experience requirements is provided in subsection 4.3.2.1.

Please see subsection 4.3.2.3 for an organizational chart for depth of talent across AI Success/Project management, AI Engineering, Data Science and extended team through the Partner Ecosystem to ensure Team DataRobot can fulfill requirements enumerated in the RFP.

4.3.1.3

Vendor should address how they will make substitutions for proposed staff if staff leave the project before completion. The plan should address Agency's prerogative to accept or reject proposed replacements for any or for no cause.

Team DataRobot will make substitutions only as necessary and will ensure any substituted staff meet original vendor requirements. The team will make reasonable efforts to ensure continuity between team members, including where possible, a proper introduction/handoff with relevant client team, as well as a reasonable transition period. DHHR will have the prerogative to accept or reject proposed replacements for any or for no cause.

4.3.2 Mandatory Qualifications/Experience Requirements

The following mandatory qualification/experience requirements must be met by the Vendor as a part of its submitted proposal. Vendor should describe how it meets the mandatory requirements and include any areas where it exceeds the mandatory requirements. Failure to comply with mandatory requirements will lead to disqualification, but areas where the mandatory requirements are exceeded will be included in technical scores where appropriate. The mandatory qualifications/experience requirements are listed below.

4.3.2.1

For each of the accounts listed in response to 4.3.2.1, vendor shall provide a reference including Name, title, email address, phone number, and the date range in which services were provided.

Customer	Reference	Description
Department of Homeland Security (DHS) - U.S. Customs and Border Patrol	Kevin Kovack IT Specialist 703-254-9832 kevin.kovack@cbp.dhs.gov Date range of work: January 2017 - present	For nearly four years, DataRobot has enabled AI applications for homeland security across 12+ mission areas across multiple DHS agencies in a wide variety of data modeling projects. For example, DataRobot's AI Cloud Platform has enabled effective and secure border transportation by predicting activity at crossing points to support better decisions about staffing levels. This use case helps reduce wait times to spur economic trade, as well as ensure enough personnel are on hand to screen for illegal goods and criminal activity. With DataRobot, the average model deployment timeline was expedited from 12 months to 3 months, where 200,000,000+ predictions are made per year supporting critical national security missions.
Food and Drug Administration (FDA)	John Wan (Supervisory Operation Analyst in Division of Quality Data Science) john.wan@fda.hhs.gov 240-753-3384 (work cell) Date range of work: December 2020 - present **Customer with DataRobot since March 2020; reference became primary POC for DataRobot in December 2020*	In March 2020, DataRobot began their support of the Food and Drug Administration (FDA). The FDA team was aiming to empower multiple user profiles to quickly and accurately build and iterate on predictive models. Over the first two months of the engagement, the DataRobot team collaborated on three data modeling projects where they held weekly training sessions with eight users. The data modeling projects focused on quickly iterating on models to more efficiently identify where inspection and enforcement resources should be sent for drug manufacturing facility inspection to keep the U.S. pharmaceutical market safe. Currently, the DataRobot team continues to meet with the FDA users in biweekly data science trainings to hone the users' proficiency on the platform through work on additional data modeling projects. Additionally, FDA is focusing on producing insightful and actionable predictions and has explored displaying the facility inspection predictions on dashboards for consumption. FDA and DataRobot also hosted an AI eminence workshop with over 90 attendees to discuss the AI/ML strategy at large within the agency.

In addition to the references, below are some additional select government accounts. The DataRobot platform is inherently use case agnostic. Customers have leveraged the platform to train models with their own data, adapting to a large variety of problem domains. Due to non-disclosure agreements, we cannot provide additional detail into the below data modeling projects in this document beyond what is provided (including supplemental

links or references). Please note that we will be able to provide more deep information. If DataRobot is selected from this RFP.

Additional Customer	Description
U.S. Army	DataRobot has worked extensively with the Army on numerous projects starting with our support of the U.S. Army Office of Business Transformations (OBT) and extending to U.S. Army Tank and Automotive Command (TACOM) and the Undersecretary of the Army for Financial Management & Comptroller (FM&C). In September 2020, OBT awarded a $1.7M contract to DataRobot to design, build, test, and deploy an Artificial Intelligence / Machine Learning (AI/ML) solution to resolve Unliquidated Financial Obligations (ULOs). The model is 3+ times more effective than the previous method, and over $2B has already been deobligated. DataRobot also worked with the Army Audit Agency (AAA) to build multiple anomaly detection models to predict specific fraud patterns and provide alerts on potentially risky behavior on the Army's Government Purchase Card (GPC). DataRobot's models reduced the number of anomaly alerts by 66% and increased the true-positive rate by 300%. DataRobot also supported an AI/ML initiative called Deep Green. This kicked off in January 2021 with the "challenge" to forecast the monthly readiness of M2A3 Bradley Fighting Vehicles. Fifteen (15) teams and over 80 participants competed. DataRobot provided both software and interactive training sessions to answer questions about preparing the dataset for modeling, feature engineering, and time series modeling capabilities in the platform. This competition enabled the Army workforce with the AI/ML skills needed to prepare data for AI/ML modeling, conduct feature engineering, and deploy working solutions. Working with Army FM&C, the DataRobot team built and deployed multiple AI/ML models to help process unmatched financial transactions (UMTs) faster and more efficiently than historic methods. When deployed into production, the DataRobot solution reduced the processing time for a single UMT from 2 hours to 2 minutes, helps process more than 30,000 PDF vouchers annually, and saves over 700,000 full-time employee (FTE) hours per year. Finally, the DataRobot team has supported Army TACOM with software and services to build, test, and deploy predictive maintenance models for non-combatant vehicles to provide Army maintainers with early identifiers of vehicles which are close to failure and in need of major repair.
State of Maryland	DataRobot has supported the State of Maryland within MDThink (Maryland's Total Human-services Integrated Network) to forecast Temporary Cash Assistance Caseloads for resource and budget planning utilizing time series modeling. Initial performance of this model reduced error of traditional forecasting method by over 200% - the team is currently working towards productionalizing the model and have developed a Qlik dashboard prototype to display predictions. Aside from support for modeling projects, DataRobot also led a strategic AI roadmapping workshop to identify and prioritize fifteen additional applications of AI/ML.
State of West Virginia Auditor's Office	The State of West Virginia Office of Auditor is mitigating fraud, misuse, and abuse in the purchasing process by using machine learning. Applying augmented intelligence to millions of yearly purchase card transactions is protecting taxpayer

	dollars and increasing organizational productivity. This is a live application being utilized on a daily basis.

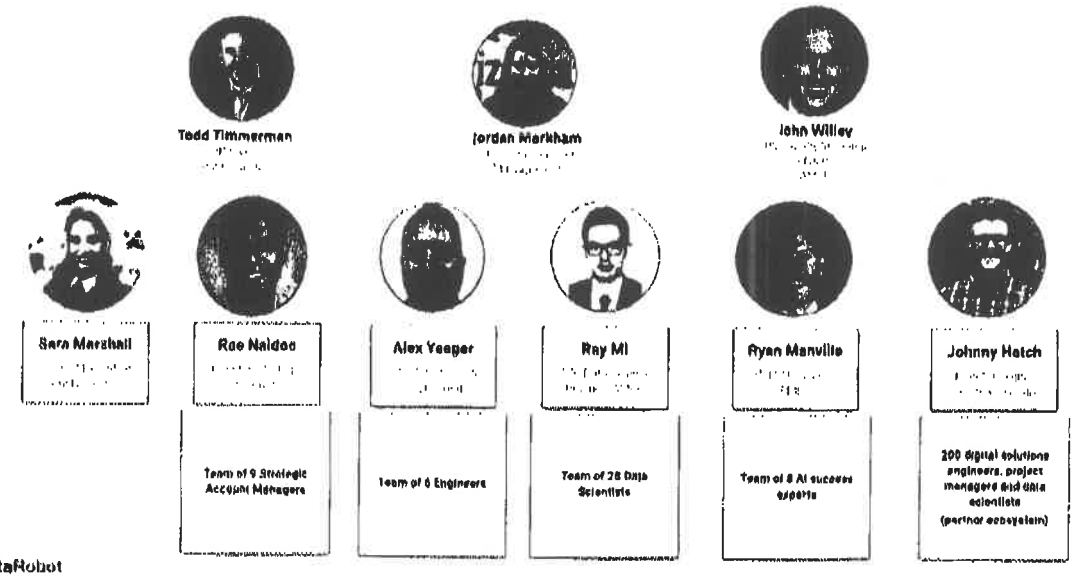
Finally, our team also has developed a deep understanding of the DHHR landscape through our engagements with ODCP and BBH over the past year. This would enable Team DataRobot to quickly re-onboard and continue making progress with the organizations. Team DataRobot is confident that we can effectively pair this critical organizational understanding of DHHR with relevant lessons learned across our federal, state/local government and commercial accounts.

4.3.2.3

Vendor must provide an organization chart with sufficient detail to demonstrate the ability to support up to the maximum number of projects over the life of the contract as well as to support the maximum number of concurrent projects at any point over the life of the contract.

Below is an organizational chart of the teams Team DataRobot can utilize on this project. These teams and numbers reflect U.S. based resources and provide fail safes and redundancies for any staffing needs that may arise to support up to 18 projects over the course of the year.

Organizational Overview



Appendix

Appendix A DataRobot Trust Package

TÜV NORD

CERTIFICATE

Management system as per

ELOT ISO/IEC 27001 : 2013

Information Technology - Security Techniques-Information Security Management Systems - Requirements

In accordance with TÜV HELLAS (TÜV NORD) S.A. procedures, it is hereby certified that

DataRobot Inc.
225, Franklin Str.
13th Floor
Boston, MA 02110
United States

with the sites according to the annexes

applies a management system in line with the above standard for the following scope

DataRobot offers an Enterprise Machine Learning Automation Platform System that empowers Users of all Skill Levels to make Better Predictions faster. Incorporating a Library of Hundreds of the most Powerful Open Source Machine Learning Algorithms, the Datarobot Platform Automates, Trains, and Evaluates Predictive Models in Parallel, Delivering more Accurate Predictions at Scale. Datarobot automates the Data Science Workflow, Enabling Users to Build and Deploy Highly Accurate Predictive Models in a Fraction of the Time of Traditional Methods. Datarobot designs, develop, maintains, and performs Support for On - Prem and Saas Solutions.

S.o.A.: Version 1.0, Dated from: 09.10.2020

Certificate Registration No. 048 20 0016
Audit Report No. IS-0178/2020

Valid from 2020-11-18
Valid until 2023-11-17
Initial certification 2020

TÜV HELLAS (TÜV NORD) S.A. Certification Body

Athens, 2020-11-18

This certification was conducted in accordance with the TÜV HELLAS (TÜV NORD) S.A. auditing and certification procedures and is subject to regular surveillance audits.

IAF ΕΣΥΔ MS Certification No of certificate 105

TÜV NORD

ANNEX

to Certificate Registration No. 048 20 0016
ELOT ISO/IEC 27001 : 2013
Information Technology - Security Techniques-Information Security Management Systems - Requirements

DataRobot Inc.
225, Franklin Str.
13th Floor
Boston, MA 02110
United States

Scope

DataRobot offers an Enterprise Machine Learning Automation Platform System that empowers Users of all Skill Levels to make Better Predictions faster. Incorporating a Library of Hundreds of the most Powerful Open Source Machine Learning Algorithms, the Datarobot Platform Automates, Trains, and Evaluates Predictive Models in Parallel, Delivering more Accurate Predictions at Scale. Datarobot automates the Data Science Workflow, Enabling Users to Build and Deploy Highly Accurate Predictive Models in a Fraction of the Time of Traditional Methods. Datarobot designs, develop, maintains, and performs Support for On - Prem and Saas Solutions.

S.o.A.: Version 1.0, Dated from: 09.10.2020

Certificate Registration No.	Site
048 20 0016-001	COLUMBUS (EASTON) 4100 Regent St, Suite S OH 43219 United States
048 20 0016-002	MINSK V. Horuzhei St., 22/1804 220123 Belarus
048 20 0016-003	SINGAPORE 5, Temasek Boulevard #11-01 Suntec Tower 5 038985 Singapore
048 20 0016-004	TOKYO Marunouchi Park Building 6F 2-6-1 Marunouchi, Chiyoda-K Tokyo 100-6906 Japan
048 20 0016-005	MUNICH Mindspace Viktualienmarkt, Rosental 7 80331 Germany

TÜV HELLAS (TÜV NORD) S.A. Certification Body

Athens, 2020-11-18

IAF ΕΣΥΔ MS Certification No of certificate 185

TÜV NORD

ANNEX

to Certificate Registration No. 048 20 0016
ELOT ISO/IEC 27001 : 2013
Information Technology - Security Techniques-Information Security Management Systems - Requirements

DataRobot Inc.
225, Franklin Str.
13th Floor
Boston, MA 02110
United States

Scope

DataRobot offers an Enterprise Machine Learning Automation Platform System that empowers Users of all Skill Levels to make Better Predictions faster. Incorporating a Library of Hundreds of the most Powerful Open Source Machine Learning Algorithms, the Datarobot Platform Automates, Trains, and Evaluates Predictive Models in Parallel, Delivering more Accurate Predictions at Scale. Datarobot automates the Data Science Workflow, Enabling Users to Build and Deploy Highly Accurate Predictive Models in a Fraction of the Time of Traditional Methods. Datarobot designs, develop, maintains, and performs Support for On - Prem and Saas Solutions.

S.o.A.: Version 1.0, Dated from: 09.10.2020

Certificate Registration No.	Site
048 20 0016-006	COPENHAGEN Raadhuspladsen 16, BC Radhuspladsen 1550 Denmark
048 20 0016-007	KHMELNYTSKYI 3rd Floor, Podilska Str. 21, BC Magnit Ukraine
048 20 0016-008	KYIV 101, Volodymyrska Str. 3rd and 4th Floors 0200 Ukraine
048 20 0016-009	KYIV Dynastia BC, 46 - 46A, Antonovycha Str., 1st Floor 02000 Ukraine
048 20 0016-010	LVIV Heroiv UPA 73, 8th Floor Lviveka Oblast Ukraine

End of the List

TÜV HELLAS (TÜV NORD) S.A. Certification Body

Athens, 2020-11-18

IAF ΕΣΥΔ MS Certification No of certificate 189

DataRobot Managed AI Cloud Security

DataRobot, the leader in augmented intelligence, is developed and built with the enterprise in mind. We implement risk-based and standards-based security protocols to secure both our services and customer data. As a part of our comprehensive security program, our managed cloud service is SOC 2 Type II and ISO 27001 certified by independent third-party auditors to ensure compliance with industry standards and best practices for information security, corporate controls, and software development.

The DataRobot platform includes a range of features and functionality that allows customers to confidently deploy our products in a variety of environments. Our engineering team has built product security, high availability, modularization, and connectivity into the application, allowing enterprises to focus on maximizing ROI through the most advanced machine learning techniques.

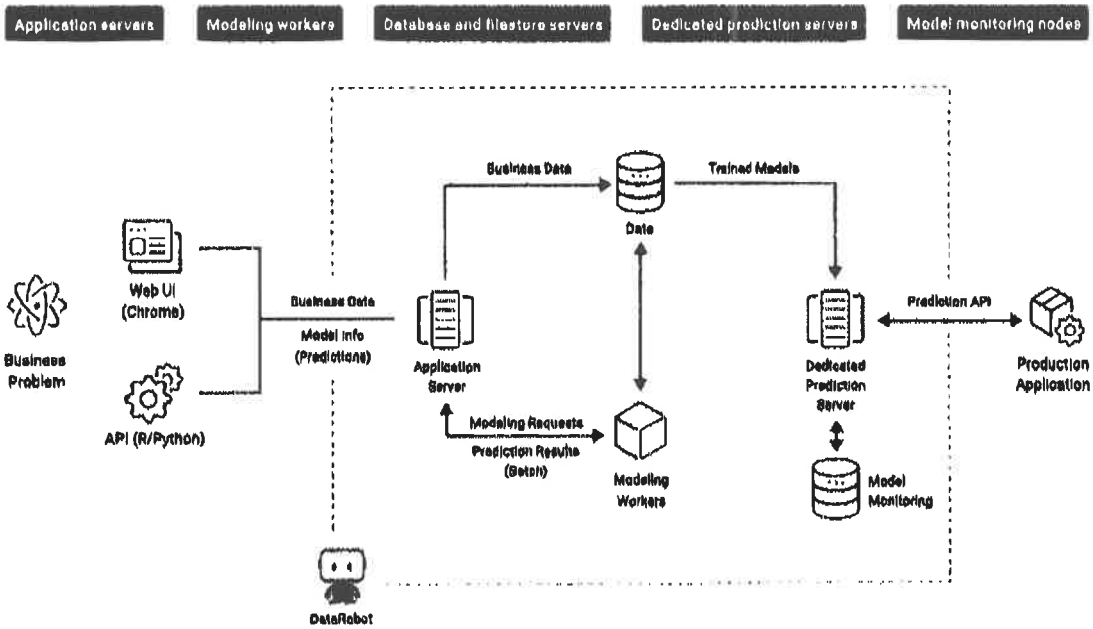
DataRobot employs many layers of security to help protect customer data. Our security programs address the following requirements:

- Access Controls
- Availability
- Processing Integrity
- Data Confidentiality and Privacy
- Physical Security

DataRobot is available as both a managed cloud service built on top of the Amazon Web Services (AWS) infrastructure, as well as an on-premise, private cloud, or hybrid cloud offering deployable on your platform of choice — AWS, Google Cloud Platform (GCP), Microsoft Azure, and virtual machines, plus bare metal, and Hadoop. This document details the security protocols we have implemented for the Managed AI Cloud offering. Please consult the DataRobot On-Premise/Private Cloud Security data sheet for security specifications for the on-premise, private cloud, and hybrid cloud offering.

Architectural Overview

The DataRobot environment is made up of the following components:



Access Control

Users interact with the external interfaces of the DataRobot application via a web-based user interface or by programming against the DataRobot APIs using DataRobot client libraries or RESTful APIs; components required for these options communicate internally within the cluster with no external exposure. Additionally, and preferably, users can generate predictions using the DataRobot Prediction Server API, described below. Any required Internet-based communications use TLS 1.2 to protect the confidentiality of the authentication process as well as the data in-flight.

WEB-BASED AUTHENTICATION

To log into the application website, users can choose to authenticate by providing a username and password or they can delegate authentication to Google. The authentication process is handled over HTTPS using TLS 1.2 to the application server. After the user sets their password, it is hashed and uniquely salted using SHA 512 and further protected with Password Based Key Derivation Function 2 (BPKDF2). The original password is discarded.

When creating passwords, only printable ASCII characters may be used. Passwords must contain at least one capital letter and one number, and be between 8 and 512 characters. The username and password cannot be the same. A user has five attempts to authenticate successfully before DataRobot locks the account, and an administrator is needed to unlock it.

DataRobot also provides enhancements to password-based authentication, including support for multifactor authentication (MFA) with software tokens generated using Time-based One-time Password (TOTP).

API AUTHENTICATION

There are two different APIs used for communicating with the DataRobot platform. All API communications use TLS 1.2 to protect the confidentiality of authentication materials.

DataRobot API. When interacting with the DataRobot API, authentication is performed using a bearer token contained in the HTTP Authorization header.

Prediction Server API. When interacting with prediction servers via the API, authentication is performed using HTTP Basic Authentication with a username and an assigned API token for the password. An additional HTTP Header named "datarobot-key" is required to further limit access to the prediction servers.

Authorization

The Managed AI Cloud service is a multi-tenant solution, but data is partitioned at the project level. That is, DataRobot stores all file data under project ID locations and manages access control on a per-project basis.

Data access is initially granted at the project level by the project creator (the owner), who can then assign specific roles and invite others to participate. Each role has a set of pre-defined capabilities, allowing careful control of who can see, modify, and delete projects. The DataRobot administrator can control user accounts, but has no access to user projects unless invited to participate by the owner.

Data Confidentiality

Data is secured at-rest using encryption. The Managed AI Cloud service uses the AWS S3 file system, which is secured with Server-Side Encryption with Amazon S3-Managed Keys (SSE-S3). On top of file system encryption, all data transferred to and from AWS S3 is encrypted in transit using TLS 1.2 for metadata stores, along with password enablement.

When using the DataRobot managed cloud environment, data sent to be trained or scored is transmitted encrypted with TLS 1.2 over the public Internet. If your data submitted to DataRobot is subject to regulatory compliance (for example, HIPAA or PII) or has a classification that requires specific security controls, you may want to consider an On-Premise AI Cluster deployment, which can be further configured to meet your compliance needs.

In addition to DataRobot's Managed AI Cloud AWS instance hosted in the US, a separate instance is hosted in the EU (Ireland) to support the unique data protection requirements of EU customers.

DATA REMOVAL

DataRobot provides two options for deleting projects. Owners can delete projects so that they are removed from active project management listings. Completely removing the project, including the data used to build it, requires an administrator. Managed AI Cloud users wanting to permanently remove data would ask their primary DataRobot contact to file a ticket with DataRobot Support and request the data be removed. Additionally, for those projects mistakenly deleted from the active list, DataRobot Support has the capability to easily restore them.

Code and Component Security

DataRobot releases new code to the Managed AI Cloud almost every week. Each release is thoroughly vetted by a formal change management process, and robust validation, testing, and enterprise-grade quality assurance programs. For example, all code commits are subjected to more than 20,000 unique tests to compare known inputs to saved expected outputs. Results are analyzed by data scientists, data engineers, and software developers with significant data science expertise. If any change causes a deviation of more than six decimal points, the code will not be accepted.

In addition to the rigorous testing for each release, DataRobot blueprints (the logic that drives the thousands of tasks that go into a machine learning project) are automatically tested every night using a series of datasets that are carefully curated to utilize every major feature, and simulate the wide variety of use cases and dataset types that DataRobot would routinely encounter in a production environment. Any regressions in key metrics are carefully reviewed by our data science team.

Source code and binary artifacts for all third-party and internal modeling libraries are stored in an internal artifact repository. New and upgraded libraries are subjected to open source license review, build compatibility testing, and extensive performance and compatibility tests. Open source components are regularly scanned for known vulnerabilities and license compliance issues, and upgraded in accordance with our validation framework to deliver consistent and reliable results, with enterprise-grade security and regulatory compliance.

Physical and Corporate Security

DataRobot is hosted on AWS, and Amazon's data centers are ISO 27001 certified, and have PCI/DSS Service Provider Level 1 and other certifications. Please refer to AWS Services in Scope for details of these certifications. AWS installs robust surveillance and detection in and around their data centers with robust perimeter security, Closed Circuit Television (CCTV) cameras, multifactor authentication mechanisms, and intrusion detection. DataRobot also has strong physical security controls in place in its offices, with badge system controlled access and intrusion detection. HR background checks are performed on all employees.

Security Monitoring and Alerting

DataRobot uses many security tools to continuously monitor the production AWS/SaaS environment, including AWS GuardDuty at the network level and Sophos Endpoint Security at the instance level.

Disaster Recovery

Disaster recovery plans are reviewed on an annual basis, with accountability and responsibility led by our Chief Operating Officer. In the case of a business continuity event, DataRobot notifies customers through proper escalation channels.

Compliance Certifications

DataRobot's Managed AI Cloud service is SOC 2 Type II and ISO 27001 certified.

DataRobot
On-Premise/Private/Hybrid AI Cloud Security

DataRobot, the leader in augmented intelligence, is developed and built with the enterprise in mind. We implement risk-based and standard-based security controls to secure both our services and customer data. As a part of our comprehensive security program, our on-premise/private cloud environment is SOC 2 Type II and ISO 27001 certified by independent third-party auditors to ensure compliance with industry standards and best practices for information security, corporate controls, and software development.

The DataRobot platform includes a range of features and functionality that allows customers to confidently deploy our products in a variety of environments. Our engineering team has built product security, high availability, modularization, and connectivity into the application, allowing enterprises to focus on maximizing ROI through the most advanced machine learning techniques.

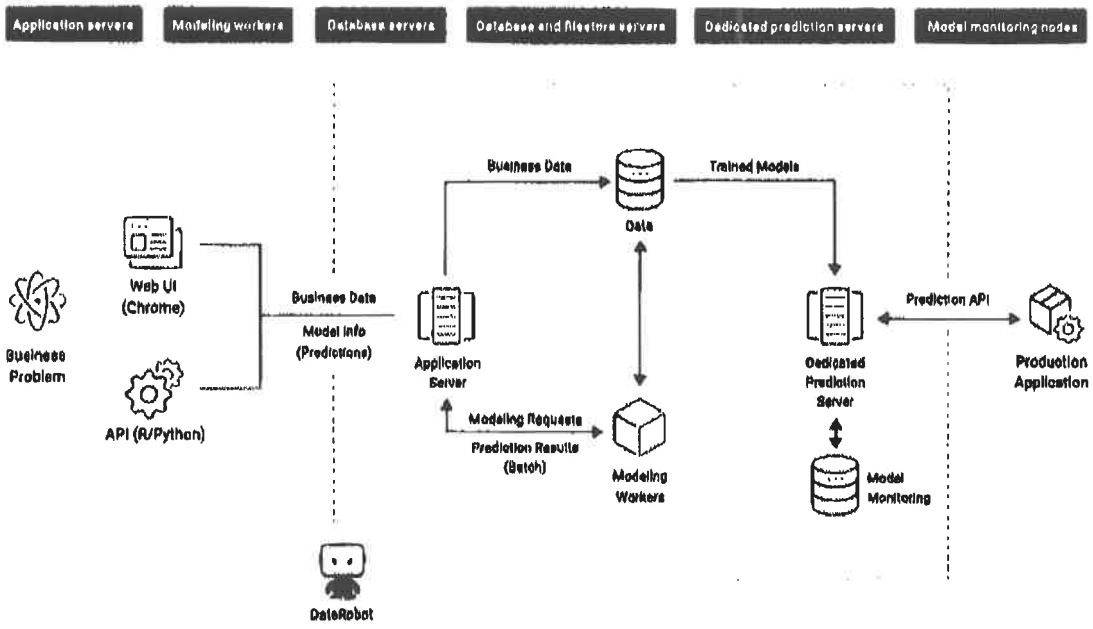
DataRobot employs many layers of security to help protect customer data. Our security programs address the following requirements:

- Access Controls
- Availability
- Processing Integrity
- Data Confidentiality and Privacy
- Physical Security

DataRobot is available as both a managed cloud service built on top of the Amazon Web Services (AWS) infrastructure, as well as an on-premise, private cloud, or hybrid cloud offering deployable on your platform of choice – AWS, Google Cloud Platform (GCP), Microsoft Azure, and virtual machines, plus bare metal, and Hadoop. This document details the security protocols we have implemented for the on-premise, private cloud, and hybrid cloud offering. Please consult the DataRobot Managed AI Cloud Security data sheet for security specifications for the managed cloud service offering.

Architectural Overview

The DataRobot environment is made up of the following components:



Access Control

Users interact with the external interfaces of the DataRobot application platform via a web based user interface or by programming against the DataRobot APIs using DataRobot client libraries or RESTful APIs; components required for these options communicate internally within the cluster with no external exposure. Additionally, and preferably, users can generate predictions using the DataRobot Prediction Server API, described below. Any required internet based communications use TLS 1.2 to protect the confidentiality of the authentication process as well as the data in-flight.

WEB-BASED AUTHENTICATION

To log into the application website, users can choose to authenticate by providing a username and password or delegate authentication to LDAP. SSO using SAML 2.0 is also supported. For Hadoop installations, Kerberos and SAML impersonation can be used. The authentication process is handled over HTTPS using TLS 1.2 to the application server. After the user sets their password, it is hashed and uniquely salted using SHA-512 and further protected with Password-Based Key Derivation Function 2 (BPKDF2). The original password is discarded.

When creating passwords only printable ASCII characters may be used. Passwords must contain at least one capital letter and one number and be between 8 and 512 characters. The username and password cannot be the same. A user has five attempts to authenticate successfully before DataRobot locks the account , and an administrator is needed to unlock it.

DataRobot also provides enhancements to password-based authentication, including support for multifactor authentication (MFA) with software tokens generated using Time-based One-time Password (TOTP).

API AUTHENTICATION

There are two different APIs used for communicating with the DataRobot platform. All API communications use TLS 1.2 to protect the confidentiality of authentication materials.

DataRobot API. When interacting with the DataRobot API, authentication is performed using a bearer token contained in the HTTP Authorization header.

Prediction Server API. When interacting with the prediction servers, authentication is performed using HTTP Basic Authentication with a username and an assigned API token used for the password.

Authorization

Data for the on premise/private cloud environment is partitioned at the project level. That is, DataRobot stores all file data under project ID locations and manages access control on a per project basis.

Data access is initially granted at the project level by the creator of the project (the owner), who can then assign specific roles and invite others to participate. Each role has a set of pre-defined capabilities, allowing careful control of who can see, modify, and delete projects. DataRobot administrators cannot access user projects unless invited to participate by the owner.

Data Confidentiality

Data is secured at-rest using the encryption of the underlying file system -- Linux, HDFS, Amazon S3, Microsoft Azure Blob Storage, or Google Cloud Storage. We recommend enabling file system encryption prior to the DataRobot installation. On top of file system encryption, all data transferred to and from the data node is encrypted in transit using TLS 1.2 for metadata stores, along with password enablement. DataRobot runs on a service account which can access your filestores, so it's important to be aware of who has access.

Data sent to be trained or scored is transmitted encrypted with TLS 1.2 over your private network to the application or prediction endpoints. If your data submitted to DataRobot is subject to regulatory compliance (for example, HIPAA, PII) or has a classification that requires specific security controls, the DataRobot platform has the ability to be installed in an off-network, bare metal setup. Offline deployments are supported with features that allow downloading scoring code (or scoring code approximations) for deployment to offline servers. Scoring through these methods is done locally and does not need to make calls over the network to a prediction server.

DATA REMOVAL

DataRobot provides two options for deleting projects. Owners can delete projects so that they are removed from active project management listings. Completely removing the project, including the data used to build it, requires an administrator. Administrators also have the ability to restore projects that users delete by accident.

Code and Component Security

DataRobot releases new code to the Managed AI Cloud almost every week. Each release is thoroughly vetted by a formal change management process, and robust validation, testing, and enterprise-grade quality assurance programs. For example, all code commits are subjected to more than 20,000 unique tests to compare known inputs to saved expected outputs. Results are analyzed by data scientists, data engineers, and software developers with significant data science expertise. If any change causes a deviation of more than six decimal points, the code will not be accepted.

Each quarter the Managed AI Cloud code branch is delivered as a separate release for on-premise, private cloud, and hybrid environments. A unique battery of tests is run using hundreds of diverse datasets to validate that all models perform as expected with every supported configuration. Testing results from the new release are compared to equivalent results from all previous releases to identify any combination of data, model type, and environment that may be experiencing a regression in accuracy, runtime, RAM usage, or any other key metric. Testing in this manner allows any potential problems or abnormalities to be quantitatively measured and proactively corrected before a release.

In addition to the rigorous testing for each release, DataRobot blueprints (the logic that drives the thousands of tasks that go into a machine learning project) are automatically tested every night using a series of datasets that are carefully curated to utilize every major feature, and simulate the wide variety of use cases and dataset types that DataRobot would routinely encounter in a production environment. Any regressions in key metrics are carefully reviewed by our data science team.

Source code and binary artifacts for all third-party and internal modeling libraries are stored in an internal artifact repository. New and upgraded libraries are subjected to open source license review, build compatibility testing, and extensive performance and compatibility tests. Open source components are regularly scanned for known vulnerabilities and license compliance issues, and upgraded in accordance with our validation framework to deliver consistent and reliable results, with enterprise-grade security and regulatory compliance.

Security Monitoring and Alerting

DataRobot provides many security tools to continuously monitor the production environment. Administrators can generate usage reports that detail individual user access, system demand, and resource allocation. All activity is also reported to the underlying syslog, allowing simple integration with your Enterprise Security Information and Event Management (SIEM) tool of choice.

Disaster Recovery

DataRobot supports replication for its storage layer. Filestore, metadata store, and real-time state information can be replicated across three different Linux servers. By making a copy of the storage layer data, you can bring DataRobot back online with the same projects and state as the source cluster.

Compliance Certifications

The DataRobot on-premise/private cloud environment SOC 2 Type II and ISO 27001 certified.

DataRobot Security Controls Report

Last Updated October 2021

Introduction

DataRobot is committed to achieving and preserving the trust of our customers through a comprehensive security and risk management strategy that impacts all aspects of the organization including our employees, systems, networks, company data, customer data, cloud environments, software development practices, and the DataRobot AI Platform. We implement industry standard solutions and best practices for security and utilize third-party audits and services to continuously assess, improve, and validate our security controls. This document provides a summary of our security controls but does not represent our security controls in their entirety.

Security Controls Summary

DataRobot's security controls are designed to protect the confidentiality, integrity, and availability of DataRobot's assets and include a number of technical, operational, and physical security controls. A summary of DataRobot's security controls includes:

Control	Summary of implemented controls
Application security	• Strong encryption of data at-rest and in-transit • 2FA and role-based access controls (RBAC) • Web Application Firewall (WAF) • Password hashing and salting

Application development security (SDLC)	• OWASP Top 10 awareness and risk mitigation • Security analysis of application code and its environment • Source code change control and feature tracking
Endpoint security controls	• Advanced malware protection • Full-disk data encryption • Secure operating system configurations
Network security controls	• Firewalls and traffic filtering • Network segmentation (VLANs) • Remote access controls (VPNs)
Cloud security controls	• AWS VPCs (Virtual Private Clouds) • AWS Security Groups and Availability Zones • AWS GuardDuty security monitoring and alerting
Corporate security governance	• Information security and acceptable use policies • Employee background checks and security training • Physical access controls (office access badges)
Compliance and certifications	• SOC 2 Type 2 certification (completed) • Other compliance initiatives that are planned or in-progress

Security Compliance & Certifications

Industry recognized compliance and security certifications are an important part of DataRobot's security strategy as they validate DataRobot's security controls through independent third-party audits. Specifically, DataRobot has achieved SOC 2 Type 2 compliance for the Trust Service Principles (TSPs) of security and confidentiality and is also working towards other compliance initiatives.

SOC2 COMPLIANCE

System and Organization Controls (SOC) compliance is governed by the Association of International Certified Professional Accountants (AICPA) and is widely considered to be an industry standard security certification for cloud or Software-as-a-Service (SaaS) solutions. A SOC audit, performed by an independent third-party auditor, is a comprehensive review of an organization's security controls for compliance to the SOC Trust Service Principles (TSPs). DataRobot has achieved the following SOC 2 certifications:

SOC 2 Certification	**TSPs**	**Received**
SOC 2 Type 1 (2018)	Security, Confidentiality	March 10, 2018
SOC 2 Type 2 (2019)	Security, Confidentiality	April 12, 2019

SOC 2 Type 2 (2020)	Security, Confidentiality, Availability, Processing Integrity	June 17, 2020
SOC 2 Type 2 (2021)	Security, Confidentiality, Availability, Processing Integrity	ETA Q4 2021

- DataRobot's SOC 2 Type 2 certification is expected to be updated annually.
- DataRobot's SOC 2 Type 2 report is available upon request (NDA required).

ISO/IEC 27001

DataRobot has achieved the following ISO/IEC certification:

ISO Certification	Description	Received
ISO/IEC 27001	Information Technology - Security Techniques - Information Security Management Systems	September 10, 2020

- The ISO/IEC 27001 certification is valid for three years.

Application Security Features

The DataRobot DataRobot AI Platform has a number of security features to protect application users, data, and traffic. The major components of the DataRobot AI Platform's security features are summarized below.

APPLICATION SECURITY FEATURES

Feature	Detail
Strong passwords	User passwords must be complex and a minimum of 8 characters.
Password encryption	User passwords are encrypted and salted using PBKDF2 (SHA512+128bit salt).
Account lockouts	User accounts are locked out after 5 failed authentication attempts.
Two-factor authentication (2FA)	2FA is available for user accounts.
API tokens	Unique API tokens for users.
Role-based access controls (RBAC)	Granular user application privilege controls are available.

Single sign-on (SSO)	DataRobot SaaS: Compatible with Google oAuth. DataRobot On-Prem: Compatible with any standard SAML 2.0 identity provider.
Web Application Firewall (WAF)	WAF is implemented to protect the DataRobot application from malicious web traffic.

APPLICATION DATA SECURITY

Feature	Detail
Data in-transit encryption	Client-server traffic is encrypted using HTTPS (TLS 1.2/AES-256).
Data at-rest encryption	Application data in AWS S3 buckets is encrypted with SSE-S3 object-level data encryption (AES-256).

Application Development Security (SDLC)

Security is a key focus in the development of the DataRobot AI Platform. Changes to the application's code are controlled and a security analysis of the DataRobot AI Platform's code and its environment is performed for every major release. The components of the Software Development Life Cycle (SDLC) security controls are summarized below.

SOURCE CODE AND FEATURE TRACKING

Control	Detail
Source code control	DataRobot uses Git to control changes to application source code - including new features, enhancements, bug fixes, and all other changes. All changes are submitted, peer-reviewed, and approved using best practices for Git.
Feature and project tracking	DataRobot uses Jira for application feature tracking, issue tracking, bug tracking, and project management functions. All features, issues, and bugs are assigned and managed using best practices for Jira.

APPLICATION CODE SECURITY ANALYSIS

Control	Detail
OWASP Top 10 security awareness	DataRobot uses the OWASP Top 10 to promote security awareness and secure coding practices for application development.
Veracode static scans	Veracode static scans are used to perform a vulnerability analysis of the application's source code to identify code-level vulnerabilities including

	vulnerabilities related to the OWASP Top 10. These scans are performed for each major DataRobot release.

- The latest Veracode report is available upon request (NDA required).

APPLICATION ENVIRONMENT SECURITY ANALYSIS

Control	Detail
Veracode dynamic analysis scans	Veracode dynamic analysis scans are used to perform a vulnerability analysis of the application's running environment to identify application-level vulnerabilities including vulnerabilities related to the OWASP Top 10. These scans are performed for each major DataRobot release.
Aquasec & Twistlock scans	Aquasec and Twistlock scans are used to perform a vulnerability analysis of the application's docker environment. These scans are performed for each major DataRobot release.

- The latest Veracode report is available upon request (NDA required).

Endpoint Security Controls

Endpoint security controls protect DataRobot endpoints from a variety of threats including malware, vulnerabilities, and other threats to data stored on endpoints in the event of compromise or theft. Endpoints include user laptops, workstations, and corporate servers. The major components of DataRobot's endpoint security controls are summarized below.

EMPLOYEE LAPTOPS

Control	Detail
Data encryption	Full-disk encryption is enabled and required on laptops (MacOS, Linux, and Windows).
Firewall	The local host firewall is enabled and required on laptops (MacOS, Linux, and Windows).
Endpoint security	An enterprise EDR (endpoint detection and response) solution is installed and required on laptops (MacOS, Linux, and Windows). The EDR solution provides protection against malware, viruses, ransomware, malicious communications, and a number of other threats.

Strong passwords	By policy, strong passwords are required for user accounts on laptops.
Patch and vulnerability management	Operating system patches, application patches, and security updates are deployed to laptops using JAMF.
Enterprise laptop management	DataRobot uses JAMF for laptop provisioning and enterprise-level laptop management features - including the deployment of standardized operating system configurations, application management, inventory, and security features.

CORPORATE SERVERS

Control	Detail
Data encryption	Full-disk encryption is enabled and required on servers that store sensitive information.
Firewall	The local firewall is enabled and required on servers that store sensitive information.
Endpoint security	An enterprise EDR (endpoint detection and response) solution is installed and required on all servers. The EDR solution provides protection against malware, viruses, ransomware, malicious communications, and a number of other threats.
Strong passwords	By policy, strong passwords are required for local accounts on servers.
Patch and vulnerability management	Operating systems and applications are required to have the latest patches and security updates installed. Patches are deployed differently for various environments.
Hardened operating systems	Best practices for security are implemented on servers to reduce their attack surface and exposure to threats. This includes limiting ports and services, restricting access with role-based access controls (RBAC), and removing unneeded services and applications.
Remote access	Remote access into any server is only permitted through the use of the authorized OpenVPN client. The VPN client enforces certificate-based user authentication and OpenSSL-256bit encryption. SSH (secure shell) is the standard protocol for server access and administration.

Network Security Controls

Network security controls protect DataRobot networks from threats such as unauthorized access, network-based attacks, and provide secure remote access into DataRobot's corporate network

environments and systems. The major components of DataRobot's network security controls are summarized below.

LOCAL AREA NETWORKS - WIRELESS

Control	Detail
Authentication	All WiFi networks enforce 802.1X (RADIUS/LDAP) authentication.
Encryption	All WiFi networks enforce WPA2-AES (AES-256) encryption.
Strong passwords	All WiFi user accounts require strong passwords.
Account lockouts	WiFi user accounts are locked out after 5 failed attempts.
Guest network isolation	The guest network is isolated from all other networks - no traffic is permitted between the guest network and other networks. The guest network enforces WPA2-PSK encryption.
Intrusion detection	Intrusion detection (IDS/IPS) policies are enabled on Access Points (APs) to detect and block malicious traffic and other network-based threats.

LOCAL AREA NETWORKS - WIRED

Control	Detail
Logical access controls	Logical access controls within the internal network environment are enforced through LDAP, role-based access controls (RBAC), firewalls, network segmentation, and other controls such as user or key-based SSH access.
Network segmentation	Network segmentation is implemented using firewalls and VLANs to control and isolate network traffic based on the network's business purpose and/or criticality. For example - business networks, engineering networks, and production networks all operate within different VLANs.
Remote access	Remote access into any corporate network is only permitted through the use of the authorized VPN client. The VPN client enforces certificate-based user authentication and OpenSSL-256bit encryption.
Boundary defense	Firewalls are deployed on all external network perimeters to protect internal networks from unauthorized access, malicious traffic, and other network-based threats.

Cloud Security Controls

Cloud security controls protect DataRobot's production cloud environment hosted within Amazon Web Services (AWS) from threats such as unauthorized access, network-based attacks, malicious

communications, and provide secure access into DataRobot's AWS accounts, instances, and services. The major components of DataRobot's cloud security controls are summarized below.

AWS INFRASTRUCTURE

Control	Detail
AWS Virtual Private Cloud (VPC)	AWS VPCs are implemented for all DataRobot AWS environments which are logically isolated from other AWS environments and provide a number of security controls such as security groups, access control lists (ACLs), and network filtering capabilities.
AWS Identity and Access Management (IAM)	AWS IAM is implemented to manage access to DataRobot's AWS services and resources securely. AWS IAM controls are used to implement role-based access controls (RBAC) for all DataRobot AWS user accounts. Strong passwords and two-factor authentication (2FA) are required for all AWS users with AWS Console access.
AWS Regions and Availability Zones	AWS Regions and Availability zones are configured and deployed for the DataRobot production environment to ensure maximum availability of the DataRobot cloud application.
AWS GuardDuty	AWS GuardDuty is enabled to provide continuous security monitoring of all AWS environments. It analyzes a variety of AWS log sources including VPC flow logs, CloudTrail event logs, and DNS logs to detect malicious traffic, malicious activity, misconfigurations, and a number of other threats within AWS environments.

AWS INSTANCES

Control	Detail
AWS Security Groups	AWS Security groups define network access control lists (ACLs) and control network access to AWS instances. They are implemented for AWS instances and act as virtual firewalls.
Access control	Secure shell (SSH) protocol is implemented for authentication and access into AWS instances. Almost all SSH authentication is configured to use key-based authentication - this includes user and service-based access.
Hardened operating systems	Best practices for security are implemented on AWS instances to reduce their attack surface and exposure to threats. This includes limiting ports and services, restricting access with key-based SSH authentication, and removing unneeded services and applications.

Endpoint / instance security	An enterprise EDR (endpoint detection and response) solution is installed and required on production instances. The EDR solution provides protection against malware, viruses, ransomware, malicious communications, and a number of other threats.

Corporate Security Governance

DataRobot has implemented a number of corporate governance policies for security including an information security policy, employee background checks, security awareness training, and physical security controls. The major components of DataRobot's corporate security governance are summarized below.

INFORMATION SECURITY POLICY

- The DataRobot Information Security Policy governs the general information security guidelines and requirements for DataRobot employees, systems, networks, and data.

ACCEPTABLE USE POLICY

- The DataRobot Acceptable Use Policy governs acceptable use of DataRobot systems, networks, and data and applies to all employees, contractors, consultants, temporary hires, and other workers.

INCIDENT RESPONSE POLICY

- The DataRobot Incident Response Policy governs the actions and procedures required to be taken by designated DataRobot employees in the event of a detected or reported security incident.

DISASTER RECOVERY POLICY

- The DataRobot Disaster Recovery Policy governs the actions and procedures required to be taken to recover from an outage or a state of degraded availability of DataRobot systems, networks, or data.

PRIVACY POLICY

- The DataRobot Privacy Policy governs how DataRobot collects and uses personal information gathered online through the DataRobot website and is available at this URL: https://www.datarobot.com/privacy/

EMPLOYEE BACKGROUND CHECKS

- Background checks are performed for all employees.
- DataRobot uses a third-party service for employee background checks and screening.

EMPLOYEE SECURITY AWARENESS TRAINING

- All employees are required to take security awareness training upon hire and then periodically on an annual basis.
- Phishing campaigns are performed monthly for all employees.
- DataRobot uses a third-party service for security awareness training.

PHYSICAL SECURITY CONTROLS

- An employee badge is required to access DataRobot offices.
- Employee badges are deactivated upon termination of employment.

Further Information

DataRobot understands the importance of providing information about the security controls implemented within DataRobot and the DataRobot Automated Machine Learning application. If further information about our security controls is needed, please contact your DataRobot account representative and we will do our best to support your request.

SIG Lite

100% Percent Complete

Tab Automation: Enable

Questionnaire Instructions:

- For each question choose either Yes, No or N/A from the drop-down menu provided. If N/A is chosen, an explanation is mandatory. Use the Additional Information Field in column F to provide.
- To display the entire contents of the tab and disable the transfer of responses from the Lite tab, select the word "Disable" in the Tab Automation field at the top of the page.
- Use the Maturity column to identify the maturity of the question. See the How To Guide for instructions on filling out this field.

Ques Num	Question/Request	Response	Maturity	Additional Information	AUP Reference	ISO 27002:2013 Relevance	
	A. Risk Assessment and Treatment						
SL.1	Is there a risk assessment program that has been approved by management, communicated to appropriate constituents and an owner to maintain and review the program?			The DataRobot Executive Management and IT & InfoSec teams assess risks on an annual basis or as deemed necessary through changes in the technological environment or industry. DataRobot uses Risk Management tools and processes to identify and prepare against business risks and threats of disruption.	A.1 IT & Infrastructure Risk Governance and Context	6.1.2	Leadership & Commitment, Information Security Risk Assessment
	B. Security Policy						
SL.2	Is there an information security policy that has been approved by management, communicated to appropriate constituents and an owner to maintain and review the policy?			The DataRobot Executive Management and IT & InfoSec teams are responsible for establishing security policies and controls for the organization. These security and confidentiality related policies are regularly reviewed by all levels within the organization. Changes to security policies must be approved by Senior Management. DataRobot's Information Security Policy, and other related policies, address the following criteria: Acceptable Use, Password Policy, Clean Desk Policy, Data Protection Policy, Mobile and Portable Storage Policy, Remote Access Policy, Security Awareness Training Policy, Network Equipment Security Policy, Server Security Policy, Operating System Security Control Policy, Amazon Web Services (AWS) Policy, Incident Response Plan Policy, Disaster Recovery Plan Policy, and Policy Compliance.	B.1 Information Security Policy Content & Maintenance	5.1.1	Policies for information security
SL.3	Have the policies been reviewed in the last 12 months?			Policies are reviewed annually at a minimum.	B.1 Procedure: d	5.1.2	Review of the policies for information security
SL.4	Is there a vendor management program?			Yes, the vendor management program is an integral part of the procurement process at DataRobot and includes participation from Legal, Enterprise Security, IT, Procurement, and Finance.			
	C. Organizational Security						
SL.5	Is there a respondent information security function responsible for security initiatives?			The DataRobot CISO is responsible for security controls, and all related security initiatives.	C.2 Security Organization Roles / Responsibilities	6.1.1	Information Security Roles and Responsibilities
SL.6	Do external parties have access to Scoped Systems and Data or processing facilities?			Please see https://www.datarobot.com/trust-center/privacy-and-personal-data/ for more details		15	Supplier relationships
	D. Asset Management						
SL.7	Is there an asset management program that has been approved by management, communicated to appropriate constituents and an owner to maintain and review the policy?			Yes and documented in policy.	D. Assessment Management	8.1	Responsibility For Assets
SL.8	Are information assets classified?			Yes based on the data inherent in the asset and whether or not it supports critical business functions.	D.1 Procedure: e.1	8.2.1	Classification of Information
	E. Human Resource Security						

SIG Lite

100% Percent Complete | Tab Automation: Enable

Questionnaire Instructions:
- For each question choose either Yes, No or N/A from the drop-down menu provided. If N/A is chosen, an explanation is mandatory. Use the Additional Information Field in column F to provide.
- To display the entire contents of the tab and disable the transfer of responses from the Lite tab, select the word "Disable" in the Tab Automation field at the top of the page.
- Use the Maturity column to identify the maturity of the question. See the How To Guide for instructions on filling out this field.

Ques Num	Question/Request	Response	Maturity	Additional Information	AUP Reference	ISO 27002:2013 Relevance	
SL_9	Are security roles and responsibilities of constituents defined and documented in accordance with the respondent's information security policy?			Yes and agreed to in the Acceptable Use Policy	C.2 Security Organization Roles/Responsibilities	6.1.1	Information security roles and responsibilities
SL_10	Is a background screening performed prior to allowing constituent access to Scoped Systems and Data?			Yes for all employees	E.3 Background Investigation Policy Content	7.1.1	Screening
SL_11	Are new hires required to sign any agreements upon hire?			Yes, Acceptable Use Policy and various employment agreements		7.1.2	Terms and conditions of employment
SL_12	Is there a security awareness training program?			Yes and includes annual training delivered quarterly	E.1 Security Awareness Training Program Maintenance	7.2.2	Information security awareness, education, and training
SL_13	Is there a disciplinarily process for non-compliance with information security policies?			Yes per the HR policy		7.2.3	Disciplinary process
SL_14	Is there a constituent termination or change of status process?			Yes per the HR policy and coordinated with IT and Enterprise Security	H.2 Revoke System and Physical Access	7.3	Termination responsibilities
	F. Physical and Environmental Security						
SL_15	Is there a physical security program?			Yes and it reports up to the CISO		5.1.1	Policies for information security
SL_16	Are reasonable physical security and environmental controls present in the building/data center that contains Scoped Systems and Data?			Yes - and these are primarily inherited from AWS	F. Physical and Environmental Security	11.1	Secure areas
SL_17	Are visitors permitted in the facility?			Yes, but AWS limits visitors		11.1.2	Physical entry controls
	G. Communications and Operations Management						
SL_18	Are Management approved operating procedures utilized?			Procedures for Operations Monitoring & Control are documented in organizational stores (Confluence/GDrive)	G. Communications and Operations Management	12.1.1	Documented Operating Procedure
SL_19	Is there an operational change management / change control policy or program that has been approved by management, communicated to appropriate constituents and an owner to maintain and review the policy?			DataRobot's change management process uses Agile processes where security related deficiencies are communicated, documented, and implemented using tools such as Confluence, JIRA, and GitHub. This Agile process is used to address security related deficiencies in the SDLC.	G.21 Change Control	12.1.2	Change Management
SL_20	Do third party vendors have access to Scoped Systems and Data? (backup vendors, service providers, equipment support maintenance, software maintenance vendors, data recovery vendors, etc.)?			Please see https://www.datarobot.com/trust/privacy-and-personal-data/ for more details		15	Supplier relationships
SL_21	Is there an anti-virus / malware policy or program (workstations, servers, mobile devices) that has been approved by management, communicated to appropriate constituents and an owner to maintain and review the policy?			Yes, and endpoint security controls include: advanced malware protection, full-disk encryption, and secure operating system (OS) configurations.	G.7 Virus Protection (Servers), G.8 Virus Protection (Workstations)	12.2.1	Controls Against Malware
SL_22	Are system backups of Scoped Systems and Data performed?				G.20 Backup Media Restoration	12.3.1	Information Back-Up
SL_23	Are firewalls in use for both internal and external connections?				G.2 Network Security – Firewall(s)	13.1.3	Segregation in networks
SL_24	Are vulnerability assessments, scans or penetration tests performed on internal or external networks?			Yes, Penetration tests are performed annually at a minimum and vulnerability scans are performed weekly.	L.2 Technical Compliance Checking – Vulnerability Testing and Remediation	12.6.1	Control of technical vulnerabilities
SL_25	Are there external network connections (Internet, intranet, extranet, etc.)?					13.1.1	Network Controls
SL_26	Is wireless networking technology used?	No		Not in our SaaS application	G.17 Unapproved Wireless Networks	13.1.1.c	Network Controls
SL_27	Is there a removable media policy or program (CDs, DVDs, tapes, disk drives) that has been approved by management, communicated to appropriate constituents, and an owner to maintain and review the policy?			N/A for our AWS hosted SaaS application	G.15 Physical Media Tracking	8.3.1	Management of Removable Media

SIG Lite

100% Percent Complete | Tab Automation: Enable

Questionnaire Instructions:

- For each question choose either Yes, No or N/A from the drop-down menu provided. If N/A is chosen, an explanation is mandatory. Use the Additional Information Field in column F to provide.
- To display the entire contents of the tab and disable the transfer of responses from the Lite tab, select the word "Disable" in the Tab Automation field at the top of the page.
- Use the Maturity column to identify the maturity of the question. See the How To Guide for instructions on filling out this field.

Ques Num	Question/Request	Response	Maturity	Additional Information	AUP Reference	ISO 27002:2013 Relevance	
SL.28	Is Scoped Data sent or received electronically or via physical media?			Yes our SaaS is a web application	G.4 Network Management – Encrypted Authentication Credentials	8.3.3	Physical media in transit
SL.29	Are Web services provided?			Our SaaS is available as a GUI and has API integrations that customers can use		N/A	
	H. Access Control						
SL.30	Are electronic systems used to transmit, process or store Scoped Systems and Data?			Yes, Okta SSO and MFA is leveraged for access to out infrastructure.		N/A	
SL.31	Are unique user IDs used for access?			Yes	H.1 Password Controls	9.2.1.a	User registration and de-registration
SL.32	Are passwords required to access systems transmitting, processing or storing Scoped Systems and Data?			Yes	H.1 Password Controls	9.4.3	Password Management System
SL.33	Is remote access permitted?			Yes, leveraging our VPN and Okta	H.7 Restrictions and Multifactor Authentication for Remote Access	6.2	Mobile devices and teleworking
	I. Information Systems Acquisition Development & Maintenance						
SL.34	Are business information systems used to transmit, process or store Scoped Systems and Data?			Databases are used within our SaaS		14.1	Security requirements of information systems
SL.35	Is application development performed?			Yes we develop and maintain our SaaS		14.2.1	Secure development policy
SL.36	Is there a formal Software Development Life Cycle (SDLC) process?			DataRobot uses an Agile process for development, JIRA for work management, and Confluence for technical documentation where any changes relevant to security controls are clearly communicated. Furthermore, GitHub is used to review, approve, and implement any application changes. Any change that impacts system security is communicated to relevant DataRobot employees and customers as a part of the Agile development and Sales & Support process.	I.7 Application Security SDLC Phases	14.2.1	Secure development policy
SL.37	Are systems and applications patched?			Yes, Criticals within 7 days, Highs within 14 days and Mediums within 180 days	I.4 System Patching	12.6.1	Management of technical vulnerabilities
SL.38	Is a web site supported, hosted or maintained that has access to Scoped Systems and Data?			The SaaS is delivered with a web site and API		N/A	
SL.39	Are vulnerability tests (internal/external) performed on all applications at least annually?				I.1 Application Vulnerability Assessments / Ethical Hacking	18.2.3	Technical compliance review
SL.40	Are encryption tools managed and maintained for Scoped Data?			AWS KMS is leveraged for our SaaS	G.22 Data Security Policy - Encryption	10.1	Cryptographic controls
	J. Incident Event and Communications Management						
SL.41	Is there an Incident Management program?			Yes, and we have a Security Incident Response Team	J.1 Information Security Incident Management P.6 Privacy Incident and Response Management	16	Information security incident management
	K. Business Resiliency					ISO 22301:2012 Relevance	
SL.42	Is there an established Business Resiliency program that has been approved by management and communicated to appropriate constituents?			Yes we have a Business Continuity Management program that has been communicated to all business stakeholders	K.1 Business Resiliency Governance	5.2	Management Commitment
SL.43	Has a Business Impact Analysis been conducted?			Yes each department has completed a business impact analysis in 2021	K.2 Business Impact Analysis	8.2.2	Business impact analysis
SL.44	Is there a formal process focused on identifying and addressing risks of disruptive incidents to the organization?			Yes, as part of the BCM Program	K.3 Risk Assessment	8.2.3	Risk assessment

SIG Lite

100% Percent Complete | Tab Automation: Enable

Questionnaire Instructions:

- For each question choose either Yes, No or N/A from the drop-down menu provided. If N/A is chosen, an explanation is mandatory. Use the Additional Information Field in column F to provide.
- To display the entire contents of the tab and disable the transfer of responses from the Lite tab, select the word "Disable" in the Tab Automation field at the top of the page.
- Use the Maturity column to identify the maturity of the question. See the How To Guide for instructions on filling out this field.

Ques Num	Question/Request	Response	Maturity	Additional Information	AUP Reference	ISO 27002:2013 Relevance	
SL-45	Are specific response and recovery strategies defined for the prioritized activities?	[redacted]		Yes with a 48 hour RTO	K.4 Response and Recovery Strategy	8.3.1	Determination and selection
SL-46	Are formal business continuity procedures developed and documented?	[redacted]		Yes in accordance with ISO 22301	K.5 Business Activity Level Recovery Planning	8.4	Establish and implement business continuity procedures
SL-47	Has senior management assigned the responsibility for the overall management of the response and recovery efforts?	[redacted]		Yes within the BCM program	K.1 Business Resiliency Governance	N/A	
SL-48	Is there a periodic (at least annual) review of your Business Resiliency Program?	[redacted]		Yes, annually at a minimum	K.7 Exercising	8.4.1	Establish and implement business continuity procedures
SL-49	Are there any dependencies on critical third party service providers?	[redacted]		AWS hosted	K.2 Business Impact Analysis	8.1 8.3 8.3.1 8.44	Operational Planning and Control Business continuity strategy Determination and selection Business continuity plans
SL-50	Is there a formal, documented exercise and testing program in place?	[redacted]		Yes as part of our BCM	K.7 Exercising	8.5	Exercising and testing
SL-51	Is there an Influenza Pandemic / Infectious Disease Outbreak Plan?	[redacted]		Yes, and we have been operating fully remotely during the ongoing pandemic with no impact to service of our SaaS	K.8 Infectious Disease Planning	N/A	
SL-52	Is there insurance coverage for business interruptions or general services interruption?	[redacted]		Yes	K.9 Business Insurance	N/A	
	L. Compliance						
SL-53	Is there an internal audit, risk management, or compliance department, or other management oversight unit with responsibility for identifying and tracking resolution of outstanding regulatory issues?	[redacted]		Yes we have a Internal Audit team within Finance and we have a Risk and Compliance team within Enterprise Security that are responsible for identifying and tracking resolution of issues	L.3 Monitoring and Reporting - Compliance Reporting L.4 Monitoring and Reporting - Compliance	18.1.1	Identification of applicable legislation and contractual requirements
SL-54	Are there policies and procedures to ensure compliance with applicable legislative, regulatory and contractual requirements to address intellectual property rights on business processes or information technology software products?	[redacted]		Yes these are covered by various legal policies and our Data Management Policy	L. Compliance	NA	
SL-55	Is there a records retention policy covering paper & electronic records, including email in support of applicable regulations, standards and contractual requirements?	[redacted]		Yes within the Information Security Policies	D.1 Asset Accounting and Inventory	18.1.3	Protection of records
SL-56	Is licensing maintained in all jurisdictions where the business is or where licensing is required?	[redacted]		Yes		NA	
SL-57	Is there an internal compliance and ethics program to ensure professional ethics and business practices are implemented?	[redacted]		Yes	A.4 Professional Ethics and Business Practices	NA	
SL-58	Are marketing or selling activities conducted directly to client's customers?	No				NA	
SL-59	Are there direct interactions with your client's customers?	No				NA	
SL-60	Are policies and procedures maintained for enabling compliance with applicable legal, regulatory, statutory, or contractual obligations related to any information security requirements?	[redacted]		Yes between the Risk and Compliance, and Legal teams	A.3 Legal, Regulatory, and Standards Compliance	NA	
SL-61	Is there a formalized governance process to identify and assess changes that could significantly affect the system of internal controls for security, confidentiality and availability?	[redacted]		Yes, we have an Enterprise Security Steering Committee that provides this governance	G.21 Change Control L.3 Procedure b.3.i	NA	
SL-62	Are accounts opened, transactions initiated or other account initiation activity applying payments, taking payments, transferring funds, etc. through either electronic, telephonic, written or in-person requests made on behalf of your client's?	No				NA	
	M. Mobile						

SIG Lite

100% Percent Complete | Tab Automation: Enable

Questionnaire Instructions:
- For each question choose either Yes, No or N/A from the drop-down menu provided. If N/A is chosen, an explanation is mandatory. Use the Additional Information Field in column F to provide.
- To display the entire contents of the tab and disable the transfer of responses from the Lite tab, select the word "Disable" in the Tab Automation field at the top of the page.
- Use the Maturity column to identify the maturity of the question. See the How To Guide for instructions on filling out this field.

Ques Num	Question/Request	Response	Maturity	Additional Information	AUP Reference	ISO 27002:2013 Relevance	
SL.63	Are constituents allowed to utilize mobile devices to within your environment? If yes, which of the following functions are allowed:	No		Not within our production environment		6.2	Mobile devices and teleworking
SL.64	View Scoped Data?						
SL.65	Process Scoped Data?						
SL.66	Delete Scoped Data?						
SL.67	Store Scoped Data?						
SL.68	Is there a mobile device management program in place that has been approved by management and communicated to appropriate constituents?			Yes and it is included in our Information Security Policy and Acceptable Use Policy			
	P. Privacy						
SL.69	Is Scoped Data transmitted, processed, or stored that can be classified as non-public information (NPI), personally identifiable information (PII), or sensitive customer financial information? If yes, describe and list types of data.			We do not allow sensitive personal data in the SaaS platform	P.1 Scoped Privacy Data Inventory and Flows	8.2.1	Classification of Information
SL.70	Is Scoped Data transmitted, processed, or stored that can be classified as protected health information, electronic health records, or personal health records? If yes, identify the classifications.	No			P.1 Scoped Privacy Data Inventory and Flows	8.2.2	Classification of Information
SL.71	For Scoped Data, is personal information about individuals transmitted to or received from countries outside the United States? If yes, identify the countries.	No			P.1 Scoped Privacy Data Inventory and Flows	N/A	
SL.72	For Scoped Data is there a dedicated person (or group) responsible for privacy compliance. If yes, describe. If no, explain reason			Our Chief Legal Officer is our Privacy Officer	P.3 Privacy Organization and Program Maintenance	16.1.1.b.3	Responsibilities and procedures
SL.73	For Scoped Data, is there a documented privacy policy or procedures to protect confidential information?			Please see https://www.datarobot.com/trustcenter/privacy-and-personal-data/ for more details	P.2 Privacy Policy and Privacy Notices	15.1.1	Information security policy for supplier relationships
SL.74	For Scoped Data are there regular privacy risk assessments conducted? If yes, provide frequency and scope. If no, explain reason.			Annually at a minimum	P.3 Privacy Organization and Program Maintenance	15.1.3.i	Information and communication technology supply chain
SL.75	Is there formal privacy awareness training for employees, contractors, and third-party users to ensure confidentiality and privacy of Scoped Data?			Yes and to meet GDPR, CCPA, and similar requirements	P.7 Privacy Awareness	7.2.2	Information security awareness, education and training
SL.76	Is there a formal process for reporting and responding to privacy complaints or privacy incidents for Scoped Data? If yes, describe. If no, explain reason.			Please see https://www.datarobot.com/trustcenter/privacy-and-personal-data/ for more details	P.8 Privacy Event Notification and Response Management	16.1.1	Responsibilities and procedures
SL.77	Is there a data classification and retention program for Scoped Data that identifies the data types that require additional management and governance?			Yes our Data Management Policy	P.1 Scoped Privacy Data Inventory and Flows	8.2	Information Classification
SL.78	Is there a documented response program to address privacy incidents, unauthorized disclosure, unauthorized access or breach of Scoped Data?			Yes	P.8 Privacy Event Notification and Response Management	16.1.1	Responsibilities and procedures
SL.79	Is Scoped Data disclosed to third parties? If yes, describe			https://www.datarobot.com/privacy/subprocessors/			
SL.80	Is Scoped Data disclosed to third parties outside of the U.S.? If yes, describe.			https://www.datarobot.com/privacy/subprocessors/			
SL.81	Are there contractual controls to ensure that Scoped Data shared with third parties is limited to defined parameters for access, use and disclosure? If yes, describe the controls, if no, explain reason.			Please see https://www.datarobot.com/trustcenter/privacy-and-personal-data/ for more details	P.4 Privacy Third Party Agreements	15.1.2	Addressing security within supplier agreements
SL.82	Is there a business associate contract in place to address obligations for the privacy and security requirements of the services provided?			Please see https://www.datarobot.com/trustcenter/privacy-and-personal-data/ for more details	P.4 Privacy Third Party Agreements	15.1.2	Addressing security within supplier agreements
SL.83	Is there a documented privacy program with administrative, technical, and physical safeguards for the protection of Scoped Data?			Please see https://www.datarobot.com/trustcenter/privacy-and-personal-data/ for more details	P.2 Privacy Policy and Privacy Notices	15.1.1	Information security policy for supplier relationships

SIG Lite

100% Percent Complete | Tab Automation: Enable

Questionnaire Instructions:
- For each question choose either Yes, No or N/A from the drop-down menu provided. If N/A is chosen, an explanation is mandatory. Use the Additional Information Field in column F to provide.
- To display the entire contents of the tab and disable the transfer of responses from the Lite tab, select the word "Disable" in the Tab Automation field at the top of the page.
- Use the Maturity column to identify the maturity of the question. See the How To Guide for instructions on filling out this field.

Ques Num	Question/Request	Response	Maturity	Additional Information	AUP Reference	ISO 27002:2013 Relevance	
SL_84	Is there a process for ensuring the accuracy of Scoped Data at the direction of the client? If yes, describe. If no, explain reason.	[redacted]		Working through the account representative	P.6 Management of Client Scoped Privacy Data	8.2.3	Handling of assets
SL_85	Is there a process to ensure that the personal information provided by an individual is limited for the purposes described in the respondent's privacy notice? If yes, describe. If no, explain reason.	[redacted]		Established privacy program	P.6 Management of Client Scoped Privacy Data	8.2.4	Handling of assets
SL_86	Are constituents regularly monitored for privacy compliance? If yes, describe. If no, explain reason.	[redacted]		Established privacy program			
SL_87	Are there documented policies, procedures, and controls to limit access based on need to know or minimum necessary for constituents? If yes, describe.	[redacted]		RBAC and need to know for all systems	P.6 Management of Client Scoped Privacy Data	9.1.1.b	Access control policy
SL_88	Are enforcement mechanisms applied to constituents who violate privacy policies or confidentiality requirements?	[redacted]		Monitoring is in place and we have a disciplinary policy for violations of Acceptable Use			
SL_89	Are transactions for covered accounts accessed, modified, or processed, including address changes and discrepancies? If yes, describe.	[redacted]		Established privacy program			
SL_90	Is customer data accessed, transmitted, processed, or stored that can be classified as consumer report information provided by a consumer reporting agency?	No					
	Q. Software Application Security						
SL_91	Is software provided?	[redacted]			Q. Software Application Security	N/A	
SL_92	Is there a secure software development lifecycle policy that has been approved by management, communicated to appropriate constituents and an owner to maintain and review the policy?	[redacted]		DataRobot uses an Agile process for development, JIRA for work management, and Confluence for technical documentation where any changes relevant to security controls are clearly communicated. Furthermore, GitHub is used to review, approve, and implement any application changes. Any change that impacts system security is communicated to relevant DataRobot employees and customers as a part of the Agile development and Sales & Support process.	Q.1 Secure SDLC Policies, Standards and Procedures	14.2.1	Secure development policy
	V. Cloud Security						
SL_93	Are Cloud Services provided? If yes, what service model and deployment model is provided (select all that apply):	[redacted]			V.1 Service and Deployment Models	4.3	Determining the scope of the information management system
SL_94	Software as a Service (SaaS)	[redacted]			V.1 Service and Deployment Models	N/A	
SL_95	Platform as a Service (PaaS)	No			V.1 Service and Deployment Models	N/A	
SL_96	Infrastructure as a Service (IaaS)	No			V.1 Service and Deployment Models	N/A	
SL_97	Private cloud	No			V.1 Service and Deployment Models	N/A	
SL_98	Public cloud	No			V.1 Service and Deployment Models	N/A	
SL_99	Community cloud	No			V.1 Service and Deployment Models	N/A	
SL_100	Hybrid cloud	No			V.1 Service and Deployment Models	N/A	
SL_101	Can clients define the legal jurisdictions where their data can be transmitted, processed or stored?	[redacted]			P.1 Scoped Privacy Data Inventory and Flows	15.1.2	Addressing security within supplier agreements

SIG Lite

100% Percent Complete

Tab Automation: Enable

Questionnaire Instructions:

- For each question choose either Yes, No or N/A from the drop-down menu provided. If N/A is chosen, an explanation is mandatory. Use the Additional Information Field in column F to provide.
- To display the entire contents of the tab and disable the transfer of responses from the Lite tab, select the word "Disable" in the Tab Automation field at the top of the page.
- Use the Maturity column to identify the maturity of the question. See the How To Guide for instructions on filling out this field.

Ques Num	Question/Request	Response	Maturity	Additional Information	AUP Reference	ISO 27002:2013 Relevance	
SL-102	Is data segmentation and separation capability between clients provided? If yes, describe.	No		Encryption is managed at the project level which is more granular than at the organizational level.	V.1 Service and Deployment Models	11.2.6, 9.1.1	Security of equipment and assets off-premises, Access Control Policy
SL-103	Is Scoped Data encrypted?			Data is encrypted in transit (TLS 1.2+) and at rest leveraging Server-Side Encryption with Amazon S3-Managed Keys (SSE-S3).	G.22 Data Security Policy - Encryption	10.1	Cryptographic Controls
SL-104	Are clients provided with the ability to generate a unique encryption key?	No			G.22 Data Security Policy - Encryption	10.1.2	Key Management
SL-105	Are clients provided with the ability to rotate their encryption key on a scheduled basis?	No			G.22 Data Security Policy - Encryption	10.1.2	Key Management
SL-106	Is standards based federated ID capability available to clients (e.g., SAML, OpenID)?			SAML 2.0	N/A	9.2.4	Management of secret authentication information of users.
SL-107	Are application self service features or an Internet accessible self-service portal available to clients? If yes, describe.	No			N/A	11.2.6, 9.1.1	Security of equipment and assets off-premises, Access Control Policy
SL-108	Is there a management approved process to ensure that image snapshots containing Scoped Data are authorized prior to being snapped?				N/A	N/A	
SL-109	Is there a cloud audit program to address client audit and assessment requirements? If yes, describe.	No			V.3 Cloud Audit Program	N/A	
SL-110	Is an agile development methodology in operation?			See above	I.2 Secure Systems Development Lifecycle Code Reviews Q.5 Secure Code Review	12.5	Control Of Operational Software
SL-111	Is there a formal process to ensure clients are notified prior to changes being made which may impact their service? If yes, describe.	No		Weekly updates are deployed	G.21 Change Control	12.1.2, 14.2.2	Change Management, System change control procedures
SL-112	Is there a scheduled maintenance window? If yes, what is the frequency?	No		Weekly updates are deployed; there is a notification banner to users in the app when we schedule and perform maintenance	V.4 Security Review of Hypervisor Configuration	12.1.2, 14.2.2	Change Management, System change control procedures
SL-113	Is there a scheduled maintenance window which results in client downtime, If yes, what is the period of the downtime?	No		Weekly updates are deployed; there is a notification banner to users in the app when we schedule and perform maintenance	V.4 Security Review of Hypervisor Configuration	12.1.2, 14.2.2	Change Management, System change control procedures
SL-114	Is there an online incident response status portal which outlines planned and unplanned outages? If yes, how long after an unplanned outage is this updated?			https://datarobot.statuspage.io/	J.1 Information Security Incident Management P.8 Privacy Incident and Response Management	12.1.2, 14.2.2	Change Management, System change control procedures
SL-115	Is there a 24x7x365 staffed phone number available to clients to report security incidents?	No		Please report all issues/incidents via our support portal at support.datarobot.com	J.1 Information Security Incident Management P.8 Privacy Incident and Response Management	12.1.2, 14.2.2	Change Management, System change control procedures
SL-116	Are applications created and released into production? If yes, what is the release frequency?			Weekly updates	C.1 Secure SDLC Policies, Standards and Procedures	12.5	Control Of Operational Software
SL-117	Is there an automated secure source code review? If yes, what is the frequency?			Prior to release	I.2 Secure Systems Development Lifecycle Code Reviews Q.5 Secure Code Review	12.5	Control Of Operational Software
SL-118	Is source code security reviewed manually? If yes, what is the frequency?			Prior to release	I.2 Secure Systems Development Lifecycle Code Reviews Q.5 Secure Code Review	12.2.1, 12.5	Control against malware, Control Of Operational Software

SIG Lite

100% Percent Complete

Tab Automation: Enable

Questionnaire Instructions:

- For each question choose either Yes, No or N/A from the drop-down menu provided. If N/A is chosen, an explanation is mandatory. Use the Additional Information Field in column F to provide.
- To display the entire contents of the tab and disable the transfer of responses from the Lite tab, select the word "Disable" in the Tab Automation field at the top of the page.
- Use the Maturity column to identify the maturity of the question. See the How To Guide for instructions on filling out this field.

Ques Num	Question/Request	Response	Maturity	Additional Information	AUP Reference	ISO 27002:2013 Relevance	
SL.119	Are automated penetration tests performed? If yes, what is the frequency?			Annually at a minimum	I.1 Application Vulnerability Assessments/Ethical Hacking	12.6	Technical vulnerability management
SL.120	Are clients provided with the ability to specify where their data will be stored? If yes, describe at what level (e.g., data center, country)?			At a regional level (e.g., US/EU)	P.6 Management of Client Target Privacy Data	8.2.3, 11.1, 15.1.2	Handling of assets, Secure areas, Addressing security within supplier agreements
SL.121	Does the ability exist to legally demonstrate sufficient data segmentation, in the event of a client subpoena or a forensics incident, so as not to impact other client's data? If using resource pooling?			Yes from the encryption at the project level	V.1 Service and Deployment Models	16.1	Management of information security incidents and improvements
SL.122	Is there a self-service portal or API call available to clients which provides the ability to place a "Legal hold" on client data which may be subject to a legal action, without impacting other clients data retention or destruction schedules?	No			P.6 Management of Client Target Privacy Data	8.2.3, 11.1, 15.1.2, 18.1.2	Handling of assets, Secure areas, Addressing security within supplier agreements, Intellectual property rights
SL.123	Is a Cloud API available to clients?			See documentation	H.3 Logical Access Authorization	N/A	
SL.124	Is there a client management portal which allows distributed business accounts (business units/departments) to be managed under a single central corporate account?			Leveraging RBAC with the primary admin account		12.1.2, 14.2.2	Change Management, System change control procedures
SL.125	Are staff required to use two factor authentication to remotely access the production cloud environment containing Scoped Data?				I.5 Application Security Vulnerability Assessments and Remediation	9.3.1, 9.4.2, 9.4.3	Use of secret authentication information, Secure log-on procedures, Password management system
SL.126	Are staff able to access client Scoped Data in an unencrypted state?	No			H.3 Logical Access Authorization	9.4.1	Information access restriction
SL.127	Are staff able to access client's encryption key?	No			H.3 Logical Access Authorization	9.4.1	Information access restriction
SL.128	Is there a process which allows the client to specifically list who from the cloud provider, will have access to their Scoped Systems and Data? If yes, describe.			Yes, the client would need to grant access to the specific user into a specific project	H.3 Logical Access Authorization	9.1.1, 9.4.1	Access Control Policy Information access restriction
SL.129	Are staff technically prevented from accessing the cloud environment via non-managed private devices?				H.3 Logical Access Authorization	9.1.1, 9.4.1	Access Control Policy Information access restriction
SL.130	Are there controls to prevent one client attempting to compromise another client in a resource pooled environment? If yes, describe.			Encryption at the project level	H.3 Logical Access Authorization	9.1.1, 9.4.1	Access Control Policy Information access restriction
SL.131	Is a default hardened base virtual image available to clients?	No		Not in our SaaS application	I.3 Secure Systems Hardening Standards	N/A	
SL.132	Can clients run their own security services within their own cloud environment? If yes, describe.			This would be a "on-premise" installation rather than a SaaS (DataRobot managed) contract	V.3 Cloud Audit Program	8.2.3, 11.1, 15.1.2, 18.1.2	Handling of assets, Secure areas, Addressing security within supplier agreements, Intellectual property rights
SL.133	Is there a specific Recovery Time Objective(s) (RTO)? If yes, specify the RTO for the scoped services.			48 hours	N/A	17.1	Information security continuity
SL.134	Are the failover sites for the underlying infrastructure running on different vendor physical systems?			Spread across multiple AWS Availability Zones that are geographically dispersed	N/A	17.1	Information security continuity
SL.135	Is the critical infrastructure running active/active at two or more sites?			Generally running across 5 Availability Zones	N/A	17.1	Information security continuity
SL.136	Are sites failed over as part of normal operation or as part of a test? If yes, what is the frequency?			During blue-green deployments that occur as frequently as weekly	N/A	17.1	Information security continuity
SL.137	Are all suppliers of critical hardware, network services and facility services involved in annual continuity and recovery tests?				N/A	17.1	Information security continuity
SL.138	Are all critical technology service providers described on an architecture diagram that includes physical systems and facilities?				N/A	17.1	Information security continuity

SIG Lite

100% Percent Complete | Tab Automation: Enable

Questionnaire Instructions:

- For each question choose either Yes, No or N/A from the drop-down menu provided. If N/A is chosen, an explanation is mandatory. Use the Additional Information Field in column F to provide.
- To display the entire contents of the tab and disable the transfer of responses from the Lite tab, select the word "Disable" in the Tab Automation field at the top of the page.
- Use the Maturity column to identify the maturity of the question. See the How To Guide for instructions on filling out this field.

Ques Num	Question/Request	Response	Maturity	Additional Information	AUP Reference	ISO 27002:2013 Reference	
SL-139	Is there sufficient redundancy capacity to ensure services are not impacted in multi-tenancy environments during peak usage and above?			Yes we leverage several AWS Availability Zones across multiple regions	N/A	17.2	Redundancies
SL-140	Do contracts include a penalty or remediation clause for breach of availability and continuity SLAs?				N/A	15.1.2, 15.2.1	Addressing security within supplier agreements, Monitoring and review of supplier services
SL-141	Is a Hypervisor used to manage systems used to transmit, process or store Scoped Data? If yes, describe the controls used to protect the hypervisor and the managed Guest Operating Systems.	No			V4 Security Review of Hypervisor Configuration	N/A	

DATAROBOT, INC.

REPORT OF CONTROLS AT A SERVICE ORGANIZATION RELEVANT TO SECURITY, AVAILABILITY, PROCESSING INTEGRITY AND CONFIDENTIALITY

THROUGHOUT THE PERIOD
OCTOBER 1, 2020 TO MARCH 31, 2021

DataRobot, Inc.
Report of Controls at a Service Organization Relevant to Security, Availability, Processing Integrity and Confidentiality
Throughout the Period October 1, 2020 to March 31, 2021

KLR

TABLE OF CONTENTS

Kahn, Litwin, Renza & Co., Ltd.
Boston • Newport • Providence • Waltham

KLR

Certified Public Accountants
and Business Consultants

I. INDEPENDENT SERVICE AUDITORS' REPORT

To the Management of
DataRobot, Inc.:

SCOPE

We have examined DataRobot, Inc.'s accompanying description of its enterprise artificial intelligence platform system found in Section IV titled "Description of DataRobot, Inc.'s Enterprise Artificial Intelligence ("AI") Platform" throughout the period October 1, 2020 to March 31, 2021 ("description") based on the criteria for a description of a service organization's system set forth in *DC 200, 2018 Description Criteria for a Description of a Service Organization's System in a SOC 2® Report* (AICPA, Description Criteria), ("description criteria") and the suitability of the design and operating effectiveness of controls stated in the description throughout the period October 1, 2020 to March 31, 2021, to provide reasonable assurance that DataRobot, Inc.'s service commitments and system requirements were achieved based on the trust services criteria relevant to Security, Availability, Processing Integrity and Confidentiality ("applicable trust services criteria") set forth in *TSP 100, 2017 Trust Services Criteria for Security, Availability, Processing Integrity, Confidentiality, and Privacy* (AICPA, Trust Services Criteria). The description indicates that certain complementary user entity controls that are suitably designed and operating effectively are necessary, along with controls at DataRobot, Inc., to achieve DataRobot, Inc.'s service commitments and system requirements based on the applicable trust services criteria. The description presents DataRobot, Inc.'s controls, the applicable trust services criteria, and the complementary user entity controls assumed in the design of DataRobot, Inc.'s controls. Our examination did not include such complementary user entity controls, and we have not evaluated the suitability of the design or operating effectiveness of such controls.

DataRobot, Inc. uses a service organization, Amazon Web Services ("AWS") to host the DataRobot, Inc. Enterprise Software as a Service ("SaaS") solution and its security monitoring tools, AWS GuardDuty, AWS CloudWatch and AWS CloudTrail; a service organization, Greenhouse, for applicant tracking system and recruiting software; a service organization, HireRight, to perform background checks and identity verification on potential employees; a service organization, Namely, Inc., to administer payroll and health benefits; a service organization, Veracode, for application and vulnerability testing associated with Enterprise AI Platform; a service organization, Zendesk, cloud-based service provider for help desk ticket support; a service organization, GitHub, a hosting service utilized for source code version control; a service organization, Grafana, for monitoring of the DataRobot, Inc. environment; a service organization, Proofpoint, for email security and protection; and a service organization, Sophos, for threat protection of the Enterprise AI Platform ("subservice organizations"). The description indicates that complementary subservice organizations' controls that are suitably designed and operating effectively are necessary, along with controls at DataRobot, Inc., to achieve DataRobot, Inc.'s service commitments and system requirements based on the applicable trust services criteria. The description presents DataRobot, Inc.'s controls, the applicable trust services criteria, and the types of complementary subservice organizations' controls assumed in the design of DataRobot, Inc.'s controls. The description does not disclose the actual controls at the subservice organizations. Our examination did not include the services provided by the subservice organizations, and we have not evaluated the suitability of the design or operating effectiveness of such complementary subservice organizations' controls.

KLR

SERVICE ORGANIZATION'S RESPONSIBILITIES

DataRobot, Inc. is responsible for its service commitments and system requirements and for designing, implementing, and operating effective controls within the system to provide reasonable assurance that DataRobot, Inc.'s service commitments and system requirements were achieved. In Section III, DataRobot, Inc. has provided the accompanying assertion titled "Management's Assertion Regarding DataRobot, Inc.'s Enterprise Artificial Intelligence ("AI") Platform" ("assertion") about the description and the suitability of design and operating effectiveness of controls stated therein. DataRobot, Inc. is also responsible for preparing the description and assertion, including the completeness, accuracy, and method of presentation of the description and assertion; providing the services covered by the description; selecting the applicable trust services criteria and stating the related controls in the description; and identifying the risks that threaten the achievement of DataRobot, Inc.'s service commitments and system requirements.

SERVICE AUDITOR'S RESPONSIBILITIES

Our responsibility is to express an opinion on the description and on the suitability of design and operating effectiveness of controls stated in the description based on our examination. Our examination was conducted in accordance with attestation standards established by the American Institute of Certified Public Accountants. Those standards require that we plan and perform our examination to obtain reasonable assurance about whether, in all material respects, the description is presented in accordance with the description criteria and the controls stated therein were suitably designed and operated effectively to provide reasonable assurance that DataRobot, Inc.'s service commitments and system requirements were achieved based on the applicable trust services criteria. We believe that the evidence we obtained is sufficient and appropriate to provide a reasonable basis for our opinion.

An examination of a description of a service organization's system and the suitability of the design and operating effectiveness of controls involves:

- obtaining an understanding of the system and the service organization's service commitments and system requirements.
- assessing the risks that the description is not presented in accordance with the description criteria and that controls were not suitably designed or did not operate effectively.
- performing procedures to obtain evidence about whether the description is presented in accordance with the description criteria.
- performing procedures to obtain evidence about whether controls stated in the description were suitably designed to provide reasonable assurance that the service organization achieved its service commitments and system requirements based on the applicable trust services criteria.
- testing the operating effectiveness of controls stated in the description to provide reasonable assurance that the service organization achieved its service commitments and system requirements based on the applicable trust services criteria.
- evaluating the overall presentation of the description.

Our examination also included performing such other procedures as we considered necessary in the circumstances.

KLR

Inherent Limitations

The description is prepared to meet the common needs of a broad range of report users and may not, therefore, include every aspect of the system that individual report users may consider important to meet their informational needs. There are inherent limitations in any system of internal control, including the possibility of human error and the circumvention of controls. Because of their nature, controls may not always operate effectively to provide reasonable assurance that DataRobot, Inc.'s service commitments and system requirements are achieved based on the applicable trust services criteria. Also, the projection to the future of any conclusions about the suitability of the design or operating effectiveness of controls is subject to the risk that controls may become inadequate because of changes in conditions or that the degree of compliance with the policies or procedures may deteriorate.

Description of Tests of Controls

The specific controls we tested, and the nature, timing, and results of those tests are listed in Section V, "Description of Tests of Controls and Results Thereof" of this report.

Opinion

In our opinion, in all material respects:

a. the description presents DataRobot, Inc.'s Enterprise AI Platform system that was designed and implemented throughout the period October 1, 2020 to March 31, 2021 in accordance with the description criteria.

b. the controls stated in the description were suitably designed throughout the period October 1, 2020 to March 31, 2021 to provide reasonable assurance that DataRobot, Inc.'s service commitments and system requirements would be achieved based on the applicable trust services criteria, if its controls operated effectively throughout that period and if the subservice organizations and user entities applied the complementary controls assumed in the design of DataRobot, Inc.'s controls throughout that period.

c. the controls stated in the description operated effectively throughout the period October 1, 2020 to March 31, 2021 to provide reasonable assurance that DataRobot, Inc.'s service commitments and system requirements were achieved based on the applicable trust services criteria, if complementary subservice organizations' controls and complementary user entity controls assumed in the design of DataRobot, Inc.'s controls operated effectively throughout that period.

Restricted Use

This report, including the "Description of Tests of Controls and Results Thereof" in Section V, is intended solely for the information and use of DataRobot, Inc.; user entities of DataRobot, Inc.'s Enterprise AI Platform system during some or all of the period October 1, 2020 to March 31, 2021, business partners of DataRobot, Inc. subject to risks arising from interactions with the enterprise machine learning automation platform system, practitioners providing services to such user entities and business partners, prospective user entities and business partners, and regulators who have sufficient knowledge and understanding of the following:

- The nature of the service provided by DataRobot, Inc.
- How DataRobot, Inc.'s system interacts with user entities, business partners, subservice organizations, and other parties.
- Internal control and its limitations.

KLR

- Complementary subservice organization controls and how those controls interact with the controls at DataRobot, Inc. to achieve the DataRobot, Inc.'s service commitments and system requirements.
- User entity responsibilities and how they may affect the user entity's ability to effectively use DataRobot, Inc.'s services.
- The applicable trust services criteria.
- The risks that may threaten the achievement of DataRobot, Inc.'s service commitments and system requirements and how controls address those risks.

This report is not intended to be, and should not be, used by anyone other than these specified parties.

Boston, Massachusetts
November 22, 2021

Kahn, Litwin, Renza & Co, Ltd.

DataRobot's Business Continuity Management Program

Our Business Continuity Management Program

DataRobot® is committed to taking reasonable steps to provide protection for essential activities should any event disrupt normal business operations. DataRobot has developed a rigorous Business Continuity Management (BCM) program to maintain the highest standards of resiliency at all times during our daily practices. BCM is an integral part of DataRobot's normal business operations.

DataRobot's BCM program considers various levels and types of disruptions that might affect a building, business district, city or a wide-scale condition within a region or multiple regions. Dedicated business continuity professionals ensure that recovery plans are documented, reviewed and tested.

DataRobot is committed to ensuring that its BCM program is comprehensive and up-to-date, particularly as new information, techniques, and technologies become available. Although we have taken significant steps to develop and implement sound business recovery plans, we cannot guarantee that systems will always be available or recoverable after a disaster or significant business disruption. However, we believe that our planning for such events is robust and consistent with industry best practices, enabling an effective response that safeguards the interests of our key stakeholders, reputation, brand and value-creating activities.

Additional Detail

DataRobot's multi-tenant cloud is hosted in Amazon Web Services (AWS). There are three production environments, each in a different AWS region. All code is stored in a cloud-hosted version control system. All production infrastructure is managed by a combination of automation applications. Automated scripts build the production code artifacts, test them, and deploy them. Most failure scenarios are handled by re-running an appropriate automation job and re-deploying cloud resources.

Services for each of the production environments are distributed across multiple availability zones within that region. Availability zones correspond to physical data centers. The failure of any individual AWS data center may cause a temporary loss of capacity, but would not affect the availability of DataRobot's service.

Database backups are created every four hours, with automated verification and validation jobs that ensure the backups can be restored successfully. In the event of a catastrophic failure the automation scripts would be used to re-deploy the infrastructure and services, and the backups would be restored into the new environment.

DataRobot is committed to taking reasonable steps to provide protection for essential activities should any event disrupt normal business operations.

Scope

DataRoot has established a Business Continuity Management (BCM) program, covering all business entities and locations, with the purpose of:

- Providing stable service to clients;
- Enabling continuity of critical functions during or immediately following a disruption.

Features of the BCM Program

- Defined global and regional governance bodies and executive ownership of BCM;
- BCM professionals responsible for creating, managing and monitoring DataRobot'spreparedness;
- A Group Business Continuity and Crisis Management Policy;
- Defined crisis management organizations and escalation protocols;
- Established crisis communications strategies for all stakeholders;
- Identification of critical activities and the planned recovery time objectives;
- Thorough risk and impact assessments of locations and processes, including critical suppliers;
- A training and awareness program for all staff that relates to their BCM roles; and
- Continued maintenance and review of arrangements to respond to emerging risks and changes to the firm.

We use the following strategies to implement our recovery objectives:

- Dedicated technical recovery facilities;
- Internal and third party work area recovery facilities;
- Remote working capabilities; and
- Local, regional and international recovery capabilities.

Further Information

For further information relating to DataRobot's business continuity and crisis management arrangements, please talk to your customer success representative.

DataRobot FAQ on International Transfer of Personal Data

What is Schrems II?

Schrems II is a case that was heard before the Court of Justice of the European Union ("CJEU") that challenged the validity of the Standard Contractual Clauses ("SCCs") to provide adequate safeguards to personal data that is transferred out of the EU. On July 16, 2020, the CJEU reaffirmed that the SCCs were a valid transfer mechanism, but said that in some cases, depending on the nature of the transfer, additional supplementary measures may be required. The European Data Protection Board ("EDBP") subsequently issued guidance describing what supplementary measures might be sufficient.[1]

What is FISA 702 and EO 12333?

Section 702 of the Foreign Intelligence Surveillance Act ("FISA 702") is a US statute that authorizes the collection, use and dissemination of electronic communications content that is stored or processed by US internet service or telecom providers by the US government for national security and foreign intelligence purposes.

Executive Order 12333 ("EO 12333") is a general directive that assigns US intelligence agencies responsibilities for different types of intelligence collection activities. Unlike FISA 702, it doesn't authorize any US government agencies to require any company to disclose data. Any such requirement must be authorized by a specific statute, such as FISA 702.

In the Schrems II decision, the CJEU found that US surveillance conducted under FISA 702 and EO 12333 didn't meet the requirement for an adequate level of protection of personal data that is required for the transfer of data under the General Data Protection Regulation ("GDPR").

How does the Schrems II ruling affect DataRobot?

DataRobot is unlikely to be targeted by FISA 702. It applies to "electronic communications service providers" which are commonly understood to be internet service and storage providers. The US government has confirmed that "most US companies do not deal in data that is of any interest to US intelligence agencies, and have no grounds to believe they do. They are not engaged in data transfers that present the type of risks to privacy that appear to have concerned the ECJ in Schrems II."[2]

---

[1] Recommendations 1/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data.
[2] Information on US Privacy Safeguards Relevant to SCCs and other EU Legal Bases for EU-US Data Transfers after Schrems II, September 2020.

In line with the US government guidance, the data processed within the DataRobot product is unlikely to be of interest to US intelligence agencies. FISA 702 is only authorized for the collection of foreign intelligence that is important to US national security.

What actions has DataRobot taken in response to Schrems II?

In response to the Schrems II ruling and the resulting EDBP guidance, we performed a review of its data transfers and the controls in place to protect our customers' personal data. We have performed transfer impact assessments for all transfers of customer data to our subprocessors and to DataRobot entities located in third countries without adequacy agreements.

What supplementary measures does DataRobot have in place to protect personal data transferred from the EU?

The EDBP identified three types of supplementary measures that enable transfer mechanisms to provide equivalent protection to what the data would receive in the EU – technical, contractual, and organizational. DataRobot has the following supplementary measures in place:

Technical Measures

Our Information Security Exhibit details the technical measures that are in place to protect customer data in our managed cloud. This includes:

- **Encryption.** All data residing in, or transiting to or from the DataRobot managed cloud product, is encrypted. Data in transit is encrypted using HTTP (TLS1.2/AES-256). Data at rest is encrypted with SSE-S3 object level data encryption (AES-256).
- **Logging and Monitoring.** We maintain logs of administrator and operator activity and data recovery events.
- **Access Management.** We monitor repeated failed login attempts, and will lock out a user account after five failed attempts[3]. Two factor authentication is available for user accounts, as well as unique user API tokens and Single Sign On (SSO) authentication.
- **Password Management.** Users must set complex passwords with length, character complexity, and non-repeatability requirements. Passwords are encrypted and salted using PBKDF2 (SHA512+128bit salt).
- **Secure Development Lifecycle.** We have policies for secure development and change control and follow industry best practices for change management. For each major release, static and dynamic analysis scans and container environment scans are performed, and any weaknesses found are remediated as deemed necessary.
- **Penetration Testing.** On an annual basis, we engage a third party to perform pen tests to detect corporate infrastructure and network vulnerabilities. Any found vulnerabilities are remediated as deemed necessary.

[3] For all DataRobot software except for Zepl.

Contractual Measures

Our Master Subscription Agreement automatically incorporates an Information Security Policy and a Data Processing Policy to protect any personal data that a customer might upload to the managed cloud. The Data Processing Policy is aligned to the Article 28 Model Clauses published by the European Commission for the transfer of data to a processor, and includes GDPR, UK GDPR, and CCPA compliant provisions. It also includes the new version of the EU Standard Contractual Clauses for transferring personal data to third countries, published on June 4, 2021, as well as the older version of the Standard Contractual Clauses adopted under Data Protection Directive 95/96/EC for transfers from the UK.

We enter into Data Protection Agreements and Standard Contractual Clauses with all subprocessors that impose substantially the same data protection obligations as those imposed on us.

Organizational

- **Data Residency.** For most DataRobot products on our managed cloud, customers located in the EU or in countries with an adequacy decision are hosted on AWS servers located in Ireland.[4]
- **Government Access Requests.** We have published a Government Data Request Policy that describes how any requests from government entities or law enforcement officials will be managed. We publish bi-annual Transparency Reports but to date have never received a government access request.
- **Internationally recognized standards.** We have achieved ISO27001 certification for information security, as well as SOC 2 Type 2 certification for the management of customer data safeguards based on security, availability, confidentiality, and privacy.

Where is customer data processed?

Customer data uploaded to the DataRobot managed cloud is hosted on AWS servers located in either the United States, or Ireland. Customers located in the EU or in countries with an adequacy decision are hosted on the Ireland installation.[5]

Customers may also send data to our Support and Success teams for help in troubleshooting issues or optimizing their AI predictive models. In that scenario, customers may use either a designated Support Portal AWS S3 bucket, Zendesk support ticket, or Box folder to provide their data to our employees. In each of those cases the data is transferred to the US. Alternatively and more preferably, customers can provide remote access to enable us to assist if necessary.

As a global company, we have employees worldwide to provide round the clock support and product maintenance to our customers. As discussed above, customers may specifically request assistance from our Support and Success teams, in which case data may also be shared with our Engineering

[4] For all DataRobot software except for Zepl and Algorithmia.
[5] For all DataRobot software except for Zepl and Algorithmia.

team to help troubleshoot. Our employees may be located in any of our subsidiary countries that are listed here.

What subprocessors does DataRobot use?

Our subprocessor list is available on our website. Customers can subscribe to receive updates when a new subprocessor is added.

Where can I get more information on DataRobot's processing of personal data?

Please visit our Trust Center to learn more about how we are working to keep our customers' data private and secured. From there, you can also request a copy of our Trust Package that contains further details and evidence of our security certifications. If you have any other questions, please contact your DataRobot representative or email privacy@datarobot.com.

DataRobot

LETTER OF ATTESTATION | DATAROBOT

December 16, 2021

Executive Summary

Pathfynder performed a Web Application Pen Test, Internal Red Team, External Penetration Test, and two Cyber Risk Assessments for DataRobot from September 20, 2021, through November 17, 2021. The primary objective of the assessment was to assess and enhance the overall security of the organization.

DataRobot performed well during these assessments, engaging with Pathfynder operators to address findings as they were discovered.

Scope of Work

ARIA Web Application Pen Test. Pathfynder conducted a Penetration Test of the Federal ARIA Web Application and Docker Containers. This penetration test was designed to identify, validate, and evaluate the security risks posed by vulnerabilities of these assets.

DataRobot Internal Red Team. Pathfynder conducted an Internal Red Team leveraging industry standard MITRE ATT&CK techniques to enumerate assets, identify vulnerabilities, and gather intelligence in an attempt to achieve access to the Crown Jewels (e.g., code repositories, cloud environments, SaaS applications) with the DataRobot organization.

DataRobot External Penetration Test. Pathfynder conducted an External Penetration Test of DataRobot's consumer-facing application and public-facing cloud assets and services. This penetration test was designed to identify, validate, and evaluate the security risks posed by vulnerabilities DataRobot's consumer and externally facing assets. The scope of this test covered app.datarobot.com, app2.datarobot.com and app.eu.datarobot.com applications.

Algorithmia & Zepl Cyber Risk Assessments. Pathfynder took a three-phased approach to assess the cybersecurity risk of recent DataRobot acquisitions to facilitate secure integration into the enterprise. The phases included an attack surface enumeration, internal threat hunt, and external penetration test of each assessed entity.

Use of this Document

This document has been prepared solely for DataRobot and its officers, directors, and employees. DataRobot shall own all rights, title, and interest in any written reports, analyses, information, or documentation prepared for DataRobot in connection with the security services provided to DataRobot. Completion of said security services does not guarantee, nor does Pathfynder for DataRobot attest that it, (i) will receive favorable results in any audits by third parties, (ii) will be safe from all information security risks, or vulnerabilities, or (iii) is in compliance with any third-party compliance program or any regulatory compliance requirements.

DocuSigned by:
Mike Mullen
4FFF421745A4406

Mike Mullen
COO, Pathfynder

Appendix B.1: Strategic Account Manager

SHAHED SERAJUDDIN
[redacted]

EDUCATION

UCLA ANDERSON SCHOOL OF MANAGEMENT — Los Angeles, CA
M.B.A., Full-Time Program, 730 GMAT — June 2019
- *Honors:* UCLA Merit Fellowship Recipient, Dean's List

BINGHAMTON UNIVERSITY — Binghamton, NY
B.S., Political Science, B.A., English Literature — May 2008
- *Honors:* Cum Laude, Dean's List
- *Entrepreneurship:* Founded apparel brand with $250K yearly revenue and 100+ retail accounts during undergraduate study

EXPERIENCE

DATAROBOT – *Enterprise AI platform valued at $6.3B (Series G)* — Chicago, IL
Director, Strategic Accounts — Aug 2022 – Present
Director, Team Lead, AI Success — Apr 2021 – Aug 2022
- Managed multi-million-dollar expansion programs as owner of $9M portfolio of customers, including 3 of top 5 largest global accounts by revenue, by overseeing relationships with C-suite executives and owning full accountability of renewal process
- Exceeded forecast by 30% for expansion of $700K+ contract for top 5 global bank by brokering renewal directly with EVP
- Launched company-wide retail industry capability as go-to subject matter expert by presenting at top commercial events (National Retail Federation), participating in webinars (LinkedIn Live), and creating content for key product launch (AI Cloud)

ZS ASSOCIATES – *Sales & marketing professional services firm with $1B+ revenue* — Los Angeles, CA
Strategy Insights & Planning Consultant — 2019 – 2021
- Steered data-driven marketing program for $1B+ product for Fortune 100 client by managing team of 16 across 6 projects including demand estimation, customer segmentation, and message testing studies
- Restructured sales force for largest business unit of $65B biotech firm by sizing target HCP population, calculating promotional response using secondary data, and redesigning sales territories, resulting in projected yearly cost savings of $25M
- Supported integration of $13B acquisition for biotech company by recommending adjacent markets for brand leadership to enter based on synergies with existing product portfolio, commercial attractiveness, and payer feasibility

NBCUNIVERSAL MEDIA, LLC — Los Angeles, CA
Strategy & Business Development MBA Intern, Fandango — 2018
- Led creation of strategic and financial plan for $11M ad sales partnership by building financial model, collaborating cross-functionally to craft business case, and presenting alongside leadership to key stakeholders, resulting in approval of plan
- Assessed entry in foreign market via partnership with $2B tech company by interfacing with firm's BoD, conducting expansive competitive analysis, and interviewing experts across region; presented recommendation to company president

MOSS CLOTHING, LLC – *Retail brand with revenue of $1.4M per year and distribution in 500+ retailers* — New York, NY
Founder, CEO — 2011 – 2016
- Oversaw sales and business development initiatives that led to distribution to 500+ retailers and media coverage from 85+ outlets, including co-branded initiatives with Coca Cola Company, Major Lazer, and Snoop Dogg
- Applied lean process methodologies to shorten order-to-cash cycle from 9 months (industry standard) to 9 weeks
- Built and managed in-house distribution center capable of shipping $170K of merchandise within 24 hours by optimizing space utilization, staffing workers, and systematizing pick and pack procedure

YOUARETV / COLLEGEONLY – *Tech platform which raised $1.2M from SoftBank, Peter Thiel, and more* — New York, NY
Co-founder, COO — 2009 – 2011
- Invented proprietary OTT streaming functionality in partnership with Adobe that enabled 100K viewers to participate in live broadcasts of game shows, talk shows and educational programming
- Created digital and offline marketing strategies that attracted 80K users within 8 months, including over 50% of the undergraduate student body at universities such as Princeton and Yale

ADDITIONAL

- *Honors:* Named one of the Top 25 Rising Stars in NY Tech by Business Insider, profiled in national Sprint campaign
- *Interests:* All things AI, Health & wellness (lost 145 lbs. as a teen), travel (completed cross-country road trip earlier this year), podcasts, reading (fav: Shantaram, Barbarians at the Gate, Shoe Dog), photography (7K followers), pug-whispering

Appendix B.2: Project Manager

[redacted]

WORK EXPERIENCE

DATAROBOT – San Francisco, California 2021 – Present
Director, AI Success
Built and managed long-term relationships with portfolio of Global 1000 clients, supporting adoption of AI/ML technologies and alignment with corporate strategy. Key skills in cross-functional collaboration, executive presentations, and strategy definition.

- Managing ~15 engagements across diverse industries, including tech, financial services, retail, and insurance. Responsible for driving and tracking customer health and adoption metrics, as well as internal and external reports on business progress.
- Ran over 20 client-facing workshops with business and technical stakeholders to generate, prioritize, and operationalize AI initiatives. Spearheaded development and implementation of said use cases, leading to ~$50M in realized value for clients.
- Led a ~10 person cross-functional team in a six-month design sprint for a Formula One constructor team to build ML-powered race prediction application. Oversaw design, QA, marketing, and product launch for thousands of public users.

BAIN & COMPANY – San Francisco, California 2018 – 2021
Manager (2020), Consultant (2018-2020)
Led management consulting teams to address clients' most critical issues in strategy, operations, and organizational design. Key skills in analytics, strategy, cross-functional presentations, stakeholder management, innovation, and leadership. Selected experience:
Vertical strategy for ~$1B division of ~$150B Fortune 25 tech company:

- Managed two direct reports and collaborated with client counterparts to prioritize solution development and develop value capture strategy based on market opportunity, competitive differentiation, customer urgency, and complexity.
- Identified ~$70B market opportunity for novel cloud software solutions across four priority industries via primary and secondary research, optimization modeling, and value capture scenario assessment.

Strategy retainer team for ~$4B device manufacturer:

- Developed roadmap to shorten product time-to-market by 35-50%, including business case for product development center of excellence, model for outsourced product testing, and processes for quality assurance.

Organizational design for global marketing team of ~$20B consumer electronics company:

- Developed org structure recommendations for ~8 key functions transitioning from regional to global operating model, taking into account industry best practices, cultural considerations, and anticipated client growth and needs.

UNILEVER – Englewood Cliffs, New Jersey 2017
Marketing Intern – Unilever Food Solutions

- Developed e-commerce strategy to grow online North American revenues by 200% over three years for Lipton and Pure Leaf tea portfolio by coordinating four functional teams; presented work to regional portfolio president.

APPLIED PREDICTIVE TECHNOLOGIES | MASTERCARD – Arlington, Virginia 2011-2016
Principal Consultant (2016), Engagement Manager (2014-2016), Business Consultant (2011-2014)
APT (acquired by Mastercard in late 2015) was a software-as-a-service company that provided business analytics capabilities and consulting services to help industry leaders optimize strategic decisions and innovate effectively. Selected experience:
Example Client Results:

- Served as lead on nine concurrent client engagements, managing ~15 employees across client, data, and product teams.
- Generated more than $4MM in annual APT contract revenues across retail, CPG, and financial services.

Example Product Management and Innovation Efforts:

- Led a six-person team on APT's first client engagement with a healthcare provider network. Developed new patient encounter benchmarking methodology, leading to client and physician reprioritization of emergency room care protocols.

Example Firm Contributions:

- Selected to lead seminar on operational efficiency at APT's annual conference, coordinating with C-level client panelists from four countries. Seminar ranked in top 5% of sessions based on attendee satisfaction.

EDUCATION

NORTHWESTERN UNIVERSITY: KELLOGG SCHOOL OF MANAGEMENT + MCCORMICK SCHOOL OF ENGINEERING – Evanston, Illinois
MMM Joint Degree - MBA with Distinction and MS in Design Innovation June 2018

- Beta Gamma Sigma – Cumulative GPA: 4.0, GMAT: 770, GRE: V169 Q170.
- Selected Leadership: Co-President, Innovation and Design Association; Academic Chair, Kellogg Board Fellows

HARVARD UNIVERSITY – Cambridge, Massachusetts
A.B. cum laude in Physics, Secondary Field in Government May 2011

- Recipient of Lowell House Elder Prize (2011) and Bok Center Certificate of Distinction in Teaching (2011)
- Selected Leadership: Harvard Glee Club President; Peer Advising Fellow; Harvard College Teaching Fellow

ADDITIONAL INFORMATION

- Awards: MMM Academic Achievement Award (2018), United States Presidential Scholar (2007), Eagle Scout (2007)

Appendix C.1: Customer Facing Data Scientist

[redacted]

Summary

Data scientist skilled in quantitative and qualitative analysis, with experience in both applied and academic research settings, and strong software development and language skills. Delivers data solutions from prototype through production

Education

University of Maryland — **College Park, MD. Jun. 2012–May 2016**
- Graduated in 2016 with an M.A. in Measurement, Statistics, & Evaluation
- Significant Coursework: Statistics, Psychometrics, Evaluation

Brown University — **Providence, RI. Sept. 2006–May 2010**
- Graduated in 2010 with an A.B. in Anthropology; graduated *magna cum laude* with honors in Anthropology and 4.0 GPA
- Significant Coursework: Linguistic, Medical, and Cultural Anthropology, Economics, Chinese Language

Data Science, Technology, and Language Skills

Data Science: Machine Learning (DataRobot, Spark MLlib, scikit-learn, H2O, TensorFlow), Big Data (Spark, Storm, Kafka, Accumulo, Hortonworks, H2O), Statistical Analysis (R), Data Manipulation and Processing (pandas, dplyr), Data Visualization (ggplot2, matplotlib, d3), Bayesian Analysis (Stan), Natural Language Processing (OpenNLP, NLTK, gensim), Generalized Linear Models, Multi-level Models, Factor Analysis, Structural Equations Modeling, Mixture Modeling

Software Development and Infrastructure: Java, Python, C#, C++, JavaScript, AngularJS, RShiny, Linux, Maven, Jenkins, Docker, Puppet, Amazon Web Services

Certifications: CompTIA Security+

Languages: French (5 yrs.), Mandarin Chinese (4 yrs.), Esperanto (3 yrs.), Middle Egyptian (2 yrs.)

Professional Experience

DataRobot, Customer Facing Data Scientist — **Washington, DC. Jul. 2019–Present**
- Guide and support DataRobot customers in framing, developing, and deploying automated Machine Learning solutions that create significant value, across public sector and commercial domains
- Advise customers at diverse levels of maturity on building up an AI-driven enterprise through consulting and training
- Educate prospects on DataRobot product, and guide them through structured Proofs of Concept to demonstrate value
- Develop sales and proposal materials for business development

Booz Allen Hamilton, Lead Scientist — **Washington, DC. Jan. 2014–Jun. 2019**
- Conducted data analysis on critical client questions using a wide range of Machine Learning, Statistics, Natural Language Processing, Big Data, and other Data Science techniques for clients including Department of Defense
- Extracted meaningful insights from both structured and unstructured data, including at scale with Big Data
- Developed, productionized, and brought to deployment a wide range of software for data transformation, data analysis, data collection, data manipulation, data visualization, simulation, and modeling
- Led and managed teams of data scientists and software engineers for complex analytic and software efforts, including task management, analytical oversight, and software development management
- Designed, constructed, operated, and maintained multiple Linux analytic clusters, including on-premises and cloud based
- Advised clients including Department of Defense on strategy and implementation for an effective Data Science enterprise
- Created engaging and informative data visualizations for a wide variety of source data, including interactive web apps
- Presented complex Data Science findings to high-level client leadership (e.g. four-star generals), with clear communication of technical concepts, limitations, and key insights
- Produced informative and rigorous technical deliverables capturing analytic results and strategic recommendations

University of Maryland Center for Advanced Study of Language, Faculty Research Assistant — **College Park, MD. Apr. 2012–Jun. 2014**
- Performed quantitative and qualitative research in culture and language for Department of Defense and Intelligence Community clients
- Developed complex algorithm, software with team, for assessing regional proficiency of Department of Defense personnel
- Assisted in qualitative analysis of Chinese language instructional techniques for Department of Defense personnel
- Collected and analyzed qualitative research data (e.g. survey) using varied statistical techniques, e.g. multiple regression
- Managed study participant recruitment, tracking, and data collection
- Contributed to experimental design and implementation

Appendix C.2: Field Engineer

ALEX YEAGER
alex.yeager@datarobot.com

TECHNICAL SKILLS:
Programming languages: Python, Bash
Operating Systems: UNIX, Linux (specifically CentOS/Red Hat)
Cloud Computing: Azure, GCP, AWS

EXPERIENCE:
DataRobot, Chicago, Illinois
Director, Success Engineering *Sept 2022 - Present*
AI Engineer *April 2020 – Sept 2022*

- Work with clients ranging from startups to Fortune 500 corporations to accelerate their deployment of machine learning models into production to achieve increased value. This includes determining the optimal method for hosting models, whether they be within the DataRobot platform or in other locations ranging from on-premise systems to managed Kubernetes clusters.
- Set up monitoring as part of MLOps to ensure that models deployed in production continue to produce value for clients or need to be refreshed with newer data.
- Develop solutions for integrating client external data sources from any location into DataRobot so that machine learning models can be trained from any data.
- Architect and deploy DataRobot's AI Cloud platform into customer environments in Google Cloud Platform Microsoft Azure, and Amazon Web Services clouds.

IPsoft Inc, Chicago, Illinois
Cognitive Solutions Engineer *October 2018 – March 2020*

- Implement IPsoft's Amelia AI platform within a large insurance company's environment, enabling front line support agents to assist their own clients faster and reduce overall call durations.
- Develop with the client new processes within the Amelia platform to further improve on Amelia's capabilities and provide further value.
- Build scripts to analyze code differences between Amelia platforms to ensure that code between environments is accurate, helping to improve platform reliability.

Senior Linux Systems Engineer *January 2016 – October 2018*

- Guided the technical side of IPsoft's relationship with a Fortune 50 client to provide further value through increased monitoring visibility into an IT environment encompassing 45,000+ virtual machines and in an 87% autonomic resolution rate of incidents to reduce potential IT service downtime and operational expenses.
- Acted as a primary point of contact between IPsoft and the client for technical account management and escalation issues.
- Built automations to augment the abilities of the IPsoft IPcenter platform to allow it to perform more complex workflow processes. Additionally, I also built automations to generate reports for clients encompassing large sets of operational data that provided insight into the performance of the client's IT environment and operations metrics while taking a fraction of the time to compile when done manually.
- Collaborated with client stakeholders to redesign and build new and highly automated change management workflow processes that reduced the amount of manual time required to get a change request reviewed and approved. This new process still provided a high level of visibility to change managers

and product owners to ensure that changes would not interfere with each other or be needlessly impactful to end users.

- Developed an Elasticsearch and Kibana POC to augment and improve on the reporting capabilities of IPcenter. Design work included building backend scripts in python to take data from the highly normalized IPcenter database and convert it into Elasticsearch-friendly documents.
- Worked with IPsoft internal teams to develop and implement solutions with existing and new technologies for a more seamless integration with the client's IT management processes. This included the proposal, development, and deployment of monitoring solutions for new client technology areas such as Dell EMC XtremIO and Cloudera CDH.

Linux/Unix Systems Engineer *July 2014 – January 2016*

- Designed and implemented diagnostic and remediation automations that would connect to Linux, Windows, and VMware servers (via powershell commands) to help reduce human intervention and ensure a faster response to any monitoring alert that arises in client IT environments.
- Monitored and maintained enterprise level Linux systems and VMware infrastructure deployments to ensure optimal uptime and performance.
- Deploy and configure Linux servers and applications as requested by clients.

Touch Support, Lafayette, Indiana
Systems Administrator *October 2010 – June 2014*

- Constructed and maintained IT solutions for clients ranging from high availability webhosting solutions to full office infrastructure.
- Maintained and improved server configurations for large-scale webhosting.
- Performed advanced, in-depth troubleshooting, rebuilding, and upgrading of webhosting servers and major system service daemons, including: Apache, MySQL, SMTP, POP/IMAP, and control panel daemons.
- Answered end-user support requests via e-mail and Kayako help desk including tickets escalated by more junior administrators.

EDUCATION:
Purdue University, *attended 2005-2013*
Coursework in Information systems, Systems Programming, Software Engineering, C, C++, Java, and Python programming.

Important: DataRobot Disclaimer

Thank you for the opportunity to submit our proposal for your consideration (the **Proposal**).

The information in our Proposal is provided for information purposes only. The information is not a warranty of any current product functionality or a promise of any future product functionality. If we mention any future product functionality please note that we may change, delay or cancel such functionality. You should not base your award of the contract on any information related to future functionality.

All information in the Proposal has been provided to the best of our knowledge. We take care to ensure that all information is correct but we cannot guarantee 100% accuracy and completeness of the information.

Our product is a standard enterprise platform and we do not warrant that it is fit for any specific purpose or use case for your business. Any decision to award the contract to us should be made on your assessment of how well our product fits your business needs.

Our Proposal is not an offer capable of acceptance. We are not able to accept any license, supply terms or other contract terms related to the supply of our product that are contained in your request for proposal. Any formal relationship between us and you and any agreement for the supply of our product is subject to good faith negotiations of a written agreement based on the DataRobot Master Subscription Agreement. Such agreement will only be binding once signed by authorized representatives of each party.

Except to the extent applicable law states that we cannot exclude or limit our liability, we do not accept any liability or any kind, howsoever arising, that you may incur as a result of your possession, reliance on and use of our Proposal.