Blue Prism ROM

Data Retention and Retrieval Policy

Document Revision 1.0

Trademarks and copyrights

Blue Prism Limited, 2 Cinnamon Park, Crab Lane, Warrington, WA2 0XP, United Kingdom
Registered in England: Reg. No. 4260035. Tel: +44 870 879 3000. Web: www.blueprism.com

Contents

Objective

This document outlines the main factors an organization should consider when establishing a protocol for a Data Retention policy for retaining information for operational or regulatory compliance needs.

Data Retention Policy

Data can accumulate quickly and dramatically, so it's important to establish a policy to define how long an organization needs to hold on to specific data. An organization should only retain data for as long as it's needed. Retaining data longer than necessary takes up storage space which could possibly lead to performance issue, instability of the digital workforce and cost increasing.

A comprehensive data retention policy outlines the business reasons for retaining specific data as well as what to do with it when targeted for disposal.

A data retention policy should treat archived data differently from backup data. An organization's backup data helps it recover in the event of data loss. Archived data is no longer actively used by the organization, but still needed for long-term retention. An organization may need data shifted to archives for future reference or for compliance. Archives are stored on cheaper storage media, so they reduce costs and the volume of primary data storage. The organization should be able to search archives easily.

What data to retain

It's critical to have a data retention policy that explains which data must be held, why and where it's being held, and for how long. Some data is required to be retained by a company's internal rules. Other data is required by law to be retained for a certain period of time. Common types of retained data include files, email messages and digital worker audit and session logs.

Blue Prism Digital Workers record all their activity in a session log that captures a detailed history of their actions on a case by case basis. This record might include - for example – every button pressed, every piece of data that is read, every decision made, etc. Where sensitive data is concerned, the logged information can be masked or omitted if necessary, in order to comply with data standards such as PCI (Payment Card Industry). The detail of the log relates directly back to the structure of the process diagram, making it easy to interpret by business users. Work items are typically captured in a Blue Prism queue. This provides Operations teams with immediate access to case status and outcomes as well as management information on case times, creation and completion dates.

Over time records and documents that are no longer used for business purposes become irrelevant but continue to take up valuable space in operational databases and computer disks. Archiving data to a cost-effective and secure tier is a best practice for freeing up space in applications and improving application and reporting performance in the organization. In some cases, information may need to be retained longer than the assigned policy if a legal hold is required to keep data available for litigation purposes or in case specific data continue to hold their value over time. Data and documents should be routinely deleted upon expiration of the retention period or expiration of the period mandated by the policy.

Compliance

Compliance is one of the major reasons for a company to retain data. It is common for an organization to establish its own data retention requirements, but in addition to that there are several laws and regulations that a company needs to consider in forming its data retention policy.

This is particularly true for organizations operating within regulated industries. For example:

- Sarbanes-Oxley Act (SOX) data retention policy for companies publicly traded in the U.S.

- Health Insurance Portability and Accountability Act (HIPAA) for healthcare organizations
- Payment Card Industry Data Security Standard (PCI DSS) for organizations that accept credit cards
- General Data Protection Regulation (GDPR), which regulates data privacy laws across the European Union. Mandates apply to personal data produced by EU citizens, whether or not the company collecting the data is in the EU, as well as any people and organizations whose data is stored within the European Union.

Blue Prism simplifies the audit process by automatically generating a detailed audit history. Blue Prism automations are self-documenting and audited automatically. Blue Prism digital worker actions are recorded by date and time. They also record any exceptions, job failures or event types that need to be reported for compliance to HIPAA, GDPR, or other standards. How long audit records should be kept depends on the organization internal policies as well as its industry compliance policies.

Archive

Blue Prism provides native archiving solutions for its digital worker session logs. Specific procedures need to be designed to archive files and email messages that are part of the automation: this aspect of data retentions is often overlooked, but emails and files pile up quickly, and some take up a lot of space, so it's important for an organization to define an archiving policy that includes documents that are not directly saved on the Blue Prism database.

Blue Prism tools for archiving include:

- System Archiving (from Blue Prism client or from Blue Prism Command line): involves the transfer of session log data from the database to a file structure, freeing up database space and allowing permanent archiving of the old data. The archived data can also be restored back to the database at any time.
- Data Gateways (version 6.5 Enterprise Edition or above): provides an easy-to-use, centralized method of pushing data out of Blue Prism for use in external systems for monitoring and reporting, long-term data storage.
- Splunk integration with Blue Prism (version 6 Enterprise Edition or above): Splunk is a data collection and analysis tool which can be used to receive Session Log data from Blue Prism runtime resources via a HTTP endpoint.

As already mentioned, external documents (such as emails and files) which are part of the automation but not directly stored in the Blue Prism database must be included in the archive strategy.

A sensible archive strategy will allow the production database and disk size to be kept to a minimum whilst still allowing retrieval of log files, should business units require them. Blue Prism provides a template document to help customers to assess and record this kind of operational task. See the Documents section of the Portal under the area: Robotic Operating Model -> Technical and Security -> Database Administration -> Archive and Backup Policy template. The Archive and Backup Policy Document gives guidance to archiving process logs.

A data retention policy should outline which media types are used for each category of data. Organizations need to focus on securing data throughout their lifecycle: the choice of where to store an archive is an important consideration to protect the integrity of the data against unauthorized access and natural disasters.

Data Retrieval

The organization should be able to search and retrieve archives easily within specified SLA's.

Routine retrieval testing: It is recommended to verify the integrity of the archives by periodically testing the recovery process to test the readability of archiving media and helps to reinforce established retrieval procedures for the RPA COE.

Index and search: It is critical to understand the business importance of indexing and searching. Retrieving needed files is crucial for compliance audits and litigation support activities. When a request for auditing is made, a company typically must provide the required data in a very short time. Retrieving important data months or years later would be problematic without a clear indexing policy.

The Blue Prism archiving process organizes the archive files automatically by date, process name and resource name making it easy to retrieve information at a later date. A similar indexing policy must be established for archiving of other documents (such as emails and files).

Security and retention: Data retrieval from an archive should be restricted to authorized personnel. Blue Prism provides roles and permission to restrict access to archiving and data retrieval functions.