IBM

Tables of Contents

IBM Cognos Analytics considerations for GDPR readiness

For PID(s): UT:30AGC

Notice:

This document is intended to help you in your preparations for GDPR readiness. It provides information about features of IBM Cognos Analytics that you can configure, and aspects of the product's use, that you should consider to help your organization with GDPR readiness. This information is not an exhaustive list, due to the many ways that clients can choose and configure features, and the large variety of ways that the product can be used in itself and with third-party applications and systems.

Clients are responsible for ensuring their own compliance with various laws and regulations, including the European Union General Data Protection Regulation. Clients are solely responsible for obtaining advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulations that may affect the clients' business and any actions the clients may need to take to comply with such laws and regulations.

The products, services, and other capabilities described herein are not suitable for all client situations and may have restricted availability. IBM does not provide legal, accounting, or auditing advice or represent or warrant that its services or products will ensure that clients are in compliance with any law or regulation.

Table of Contents

GDPR Overview

General Data Protection Regulation (GDPR) has been adopted by the European Union ("EU") and applies from May 25, 2018.

Why is GDPR important?

GDPR establishes a stronger data protection regulatory framework for processing of personal data of individuals. GDPR brings:

- New and enhanced rights for individuals
- Widened definition of personal data
- New obligations for processors
- Potential for significant financial penalties for non-compliance
- Compulsory data breach notification

Read more about GDPR

EU GDPR Information Portal: https://www.eugdpr.org/
IBM GDPR website: http://ibm.com/GDPR

Product Configuration - considerations for GDPR Readiness

The following sections provide considerations for configuring IBM Cognos Analytics to help your organization with GDPR requirements.

Data Life Cycle

There are two types of personal data that can be processed by IBM Cognos Analytics:

- Data that is gathered by the client
- Data associated with the users of the product

The data gathered by clients can come from corporate data warehouses, operational systems, acquired 3rd party or publicly available external data, and personal spreadsheets, and other common sources of data.

The sources of data consumed by IBM Cognos Analytics can contain personal data such as names, addresses, government identification numbers, internet addresses, email addresses and so on.

When a person uses IBM Cognos Analytics, certain pieces of information are retained about them. Depending upon how they are authenticated, this information can include their email address. The product can be configured to transfer personal information stored in the authentication source (for example, LDAP) to IBM Cognos Analytics. In addition, a user of IBM Cognos Analytics usually customizes the product by configuring their account with locale, time zone, email address, and so on.

There are two ways in which information about a user of IBM Cognos Analytics can be gathered:

- Product preferences entered by the user into their user profile, such as locale, time zone, and so on.
- Information that is transferred from an authentication source. For example, Active Directory and assigned to a user's profile.

The information that is entered by a user can be edited by that user. The information transferred from an authentication source is not editable by the user. This transferred information can represent sensitive personal information.

The other source of personal information is data that the user brings in to IBM Cognos Analytics. This can come from a variety of data sources – corporate data warehouses, data lakes, Microsoft Excel spreadsheets, and so on. The original acquisition of the data, notably any sensitive personal information, is outside of realm of IBM Cognos Analytics. You must understand these sources of data so that you can manage GDPR requests prior to using the data in IBM Cognos Analytics.

It is very important as a Controller to understand where sensitive personal information is stored in the various data sources processed by IBM Cognos Analytics so that you can respond to GDPR-related requests. IBM Cognos Analytics has no knowledge of what the data represents, including whether the data contains sensitive personal information.

In IBM Cognos Analytics, data is often modelled in either Framework Manager models or web-based data modules. These metadata representations are the basis for most of the analytic content produced by IBM Cognos Analytics. In other cases, content is created directly from uploaded files or data sets extracted from other data sources.

Personal information can be contained in the output produced from analytic content. For example, PDF report output. However, it is the references to sensitive personal information in the analytic content in IBM Cognos Analytics which must be managed with respect to GDPR.

IBM Cognos Analytics cannot identify a data item (for example a column in a table) as containing sensitive personal information. However, it is possible to search for specific data values in the analytic content, either programmatically or manually, and either modify or delete those values from the content. Business processes and controls can help to constrain the content which must be searched for sensitive information about one or more individuals.

IBM Cognos Analytics stores its base maps and shapes (polygons) on a cloud-hosted MapBox server. When a customer wants to visualize their data on a geospatial map, IBM Cognos Analytics calls these base maps and polygons from the MapBox server. It's important to note that no customer data is sent to the MapBox server and that Cognos Analytics restricts its requests to certain MapIDs and Polygon Keys from the Mapbox server.

Data Collection

As described previously, an IBM Cognos Analytics user can edit the following personal information. This information is used to configure the product behavior:

- User name
- Password
- Email address
- Locale
- Time zone
- Language
- Accessibility features enabled

Data Storage

IBM Cognos Analytics stores analytic content, including report outputs, datasets, and upload files in a content store contained in a relational database. The report outputs, datasets, and upload files can also be stored in an external object store in a specified location on disk.

Datasets and uploaded files are stored in an encrypted format. Do not modify this setting.

With respect to GDPR, IBM Cognos Analytics focuses on the data contained in the analytic content (for example, report or dashboard specification) or the data containers such as uploaded files and datasets.

The content store also contains previously described user data.

Because analytic content is stored in clear text and can contain personal information, you must apply access control to the data that is stored in the content store. The level of access control is up to you. Consider the following things:

- Physical access to the content store and its backups
  Who has physical access to the devices that contain the content store and its backups? Are the drives portable? If so, what controls are in place to track and monitor their location?

- Does everyone who has physical access to the content store or its backups have a legitimate reason for having access? How is this tracked and how is the access tracked?

- Where are backups stored? How long are they maintained? How are backups destroyed so that data is no longer present?

- Logical access to the content store and its backups
  Who has logical access to the content store or its backups? Is there a legitimate business reason for having access? Is access restricted to the functions necessary for an individual's business needs?
- Encryption of the content and storage of encryption keys
  Is the content store database encrypted? If not, you may want to consider encrypting it. The same is true for the external object store. Can the disk containing the external store be encrypted?
- Where are the encryption keys stored? Are they stored separately from the objects they are encrypting? Are keys stored in an external key store?

Data Access

Separation of duties

There are various roles which people assume when working with IBM Cognos Analytics. You should know the activities people in these roles need to perform and the access they can have to data, including personal data. You should limit access to what is necessary for each role, and consider that some individuals can have more than one role.

Role: Installation and configuration

In an organization, the installation and configuration of operating systems and application software may be assigned to a specific set of individuals. People in this role are not usually authorized to use the applications after they are configured. They are also usually responsible for upgrading applications to new versions.

Role: Operating system administration

The operating system (OS) administrator is responsible for the configuration and management of the operating system upon which IBM Cognos Analytics is hosted. This person should have limited, or no access to analytic content.

Commands that are run by an administrator should be logged and stored in log files. The OS administrator should be restricted to read access to these logs.

Role: Database administration

A database administrator is responsible for the creation, maintenance, and performance of databases for use in an organization. People in this role typically have access to all databases and the content contained in them but will have limited access rights in the OS.

Role: Operations

The operations role encompasses many different aspects, and in larger organizations, can be further broken down into more refined roles. Operations includes things such as:

- Backup and restore
- Application maintenance (tuning, stopping, starting, and so on)

A person in operations can have limited access rights in an application to allow them to fulfill the requirements of their position.

Role: IBM Cognos Analytics administrator

An IBM Cognos Analytics administrator has access to the content in IBM Cognos Analytics and is responsible for managing the system and its content. The person in this role has access to all analytic content, but they do not necessarily have access to the data in corporate databases. If output is not saved to the content store, it is possible that they are constrained from viewing the data being processed by the system.

Role: IBM Cognos Analytics modeler

A modeller is provided access to a variety of data sources in an organization, often to both test and production data sources for testing and troubleshooting scenarios. They often have read access, but not write access to the data.

Role: IBM Cognos Analytics eport author

A report author often has the same level of access to the same data as a modeller.

Role: IBM Cognos Analytics content consumer

A report consumer has limited access to data to which they have been granted access in the database and in the analytic content, for example, Framework Manager models. They are often limited to read access to the content to which they need access.

Role: Support

People in a support role in an organization have access to logs, analytic content, and data so that they can replicate and address issues. Support can be internal to an organization, but can also include support from vendors, including IBM support for IBM Cognos Analytics.

IBM Cognos Analytics provides extensive security capabilities that can be used to control access to data and analytic content. Users should be restricted to the capabilities appropriate for their role in the organization and granted access, such as read and write, to the analytic content they need to use to do their work.

For example, a person who primarily consumes pre-authored content may not be granted the ability to do modelling or upload data, and restricted to read and execute access to the pre-authored analytic content. A security administrator should have read access to a broader range of content since their role involves granting access to content across IBM Cognos Analytics.

This level of security helps to track that individuals are performing actions, and accessing and creating analytic content that is appropriate for their roles. This can help reduce the creation of uncontrolled content in the organization, which in turn can help control the use of and access to, personal data.

Activity logs

Most of the activities carried out by the people in the roles described previously can be logged. Logging these operations provides a record of the activities that have occurred. The activities can include ones that were requested by a user and questionable operations that might be found during a forensic audit.

It is important that the activity logs track nearly all activities, and that the logs are not editable, so that users can't change the logs.

Data Processing

By default, IBM Cognos Analytics is configured to encrypt data in motion. Do not change this.

With respect to data at rest, IBM Cognos Analytics encrypts all sensitive personal information that is stored in the content store. Uploaded files, datasets, and any temporary data files stored on disk are encrypted by default. Do not change this.

IBM Cognos Analytics includes a database for use as the content store for the analytic content it produces. If you choose to use your own database for the content store, you should encrypt it and its backups to control access to the data.

If an external object store is used to reduce the size of the content store in the database, you should consider encrypting the disk containing the object store.

You should store encryption keys in a separate location from the content being encrypted, possibly in external key stores.

Data Deletion

Requests to be forgotten in IBM Cognos Analytics Beyond the IBM Cognos Analytics environment, you must identify all personally identifiable information in an organization in the various data containers (databases) where it resides. It is important to be able to identify the data items (columns within a relational database) that contain personally identifiable information (PII) and what type of information they contain. This information that is valuable when you respond to GDPR related user requests.

In IBM Cognos Analytics, the main GDPR user request you will need to respond to is the right to be forgotten. An individual's PII can reside in the following objects in IBM Cognos Analytics:

- Framework Manager models and packages
- Reports and analyses
- Global parameters
- Saved prompt values
- Datasets
- Uploaded files
- Saved report outputs
- PowerCubes and Dynamic Cubes
- IBM Cognos Analytics users

When you respond to a request to be forgotten, you must identify the content that contains data associated with the individual and then either update or delete that content. You must also update the user's profile and all access control lists (ACLs) in which they are referenced to remove all traces of the individual.

During the development of IBM Cognos Analytics content, you can reduce the impact of requests to be forgotten. Train people to not use hard coded, uniquely identifying PII in filters, prompts, or parameters. You should review and audit analytic content so this policy is followed.

Framework Manager Models and Packages

Framework Manager (FM) models can contain a person's uniquely identifying PII as a scalar constant value. For example, name or social insurance number. When an individual requests to be forgotten, these static values must be removed from the models.

Avoid using uniquely identifying PII in filters in models, reports, and dashboards. Using prompts built upon data items that represent uniquely identifying PII should also be avoided.

When you receive a request to be forgotten, you must search either the model files or the packages in the IBM Cognos Analytics environments. If a search is performed of packages, you must have a mapping of the model file that forms the basis for the packages.

With the list of affected FM models, you must remove all occurrences of the individual's PII from the filters that contain the PII associated with the request to be forgotten. Then you must republish the packages to the affected IBM Cognos Analytics environments.

Reports and analyses

Reports, Query Studio reports, and Analysis Studio analyses should not contain filters based on PII data items. Having processes to document and audit to track this eliminates the requirement to search reports, global parameters, and saved prompt values for PII values when you received a request to be forgotten.

However, if it's possible that a report contains filters based on PII data items, you must search all reports for occurrences of a hard-coded scalar filter values associated with the individual who has requested to be forgotten. If one is found, you must either remove the filter entirely or replace the scalar value with another that does not identify the individual.

Global Parameters

A global parameter can be created to identify values of a PII data item. This practice should be avoided and you should have processes to enforce this.

If it is not possible to be sure that at least one global parameter does not contain PII values, then it is necessary to search the values assigned to global parameters so that PII values associated with the individual do not exist. If they exist, then those values must either be modified or deleted from the global parameter's list of values.

Saved Prompt Values

Like global parameters, if a report contains prompts that use PII data items, then it is likely that users will create saved prompt values. If this happens, you must search all saved prompt values. If an individual's PII is found, you must modify or delete the saved prompt value.

Datasets

Like uploaded files, data sets can contain PII. To respond to a request to be forgotten, you should identify the files associated with the individual and then to update or delete those data sets. The alternative is to have scheduled refreshes on your datasets. You can then update the source data which will update the dataset within the timeframe required by GDPR.

Uploaded Files

PPII can be in uploaded files so it is important to have a process that tracks the users who upload files, the data, and types of data that is in the files. You must also control the users that have data upload capabilities and govern the type of data users can upload. These controls help you to limit your search to users and file types when responding to a request to be forgotten. Additionally, we suggest having local operational processes that provide guidance on the types of data that can and cannot be included in uploaded files.

When responding to a request to be forgotten, you should identify the files associated with the individual and then update the files with the fields removed. Alternatively, you can delete the files that contain the associated data.

Saved report outputs

Historical data does have to be cleared of PII for legal and compliance reasons. However, if you receive a request to be forgotten, it is your responsibility to ensure that the PII is not included in future report outputs. If the PII are items queried from the database, you should update the database and rerun the report.

PowerCubes and Dynamic Cubes

When modeling and building PowerCubes with Transformer or Dynamic Cubes with Cube Designer, it is possible for a model to contain references to a category or member that identifies an individual. It is also possible that that data exists in multiple cubes that were built with those models.

Dimensions that contain PII in dimensional models should be identified and tracked outside of the model itself. For example, user social insurance number and name. You should also record the categories and members that are directly referenced in the model. When you receive a request to be forgotten, you must edit the models to remove specific category and member references. You must then rebuild, update, and redeploy all associated PowerCubes. You must republish all associated Dynamic Cubes models and restart or refresh the cubes.

IBM Cognos Analytics Users

IBM Cognos Analytics users have the same right to be forgotten. PII related to an individual can be found in the following objects in IBM Cognos Analytics:

- User profile
- Access control list (ACL)

When an individual asks to be forgotten, delete the person's IBM Cognos Analytics account. Deleting the IBM Cognos Analytics account removes the user account profile and their entire "My Content" area.

After their account has been deleted, it is still possible for a user account to be assigned to one or more objects in the content store. Depending on the method used to authenticate users, these entries could contain PII such as the user email address. Modify the user's ACL entries by doing the following things:

- If the user is the owner, assign ownership to another user
- If the user is not the owner, remove the user from the ACL or assign the access rights to another user

Data Monitoring

If you have tracked the reports, dashboards, and stories process, and display personal data, you can view the activity by running reports that are authored for the IBM Cognos Analytics audit database.

Ad hoc or self-service analytic content is difficult to track as there are no lineage mechanisms in IBM Cognos Analytics to identify content that is created for a metadata package or data module. In responding to data subject rights, be aware of the PII stored in the data processed by, and of the users of, IBM Cognos Analytics.

Right to Access

A person's data, as mentioned earlier, is usually stored in data sources outside of IBM Cognos Analytics, so the "right to access" involves identifying where that data resides and providing this data to the person exercising their right to access. IBM Cognos Analytics can be used to provide the data in a specific output format. For example, PDF.

An IBM Cognos Analytics user has access to their profile information through the product interface. You can also get information from an external authentication provider and give it to the user.

Responding to Data Subject Rights

Right to Modify

A person's data is usually stored in data sources outside of IBM Cognos Analytics. IBM Cognos Analytics is a read-only product (it does not write back to the underlying data sources) so you must update the data in the external data sources. Data can be uploaded from Microsoft Excel or CSV files, or datasets created from any of the accessible data sources, so you must ensure that the data created in IBM Cognos Analytics is synchronized with changes to the underlying data sources.

An IBM Cognos Analytics user can modify their personal profile information through the product interface. You are responsible for modifying other information stored in an authentication provider.

Right to Restrict Processing

To restrict processing of an individual's data, you must remove data from the source data sources, apply security to restrict access to the data, or anonymize the data. You can apply restrictions in IBM Cognos Analytics based on filter values (For example, government identification number), but this logic must be applied to all applications. It is simpler to apply restrictions in the source data systems.

Since data can be uploaded from Microsoft Excel or CSV files, or datasets created from any of the accessible data sources, you must ensure that the data created in IBM Cognos Analytics is synchronized with changes to the underlying data sources.

You can also disable an IBM Cognos Analytics user account to stop processing in the account.

Right to Object

See "Right to Restrict Processing".

Right to be Forgotten

As described previously, it is your responsibility to remove a person's data from an IBM Cognos Analytics instance. The more controls and processes that you have your organization, the more focused the operations can be to locate and either replace or delete a person's data.

IBM Cognos Analytics users can be forgotten by deleting their user profile and removing them from all ACLs they are associated with.

Right to Data Portability

You can use IBM Cognos Analytics to merge a person's data into a report output.