tyGraph Compliance White Paper

A comprehensive guide on our security options and best practices.

To save a copy, please choose to save as PDF so that URLs work.

tyGraph Compliance White Paper
Addressing GDPR Requirements

Updated on: August 2022

Introduction

This white paper was authored to help customers, and prospective customers understand the data security and data processing factors related to tyGraph. Given the heightened state of cybersecurity and legislation like the European Union's GDPR legislation, tyGraph must collaborate with its customers to ensure that security and data compliance requirements are met to the customer's satisfaction.

Many of our customers are global, with different legal entities registered in various jurisdictions. We also have customers with head offices on five continents. In addition, tyGraph is also a downstream analytics system to Office 365. This means that tyGraph has no way of knowing where the data is coming from or what information is contained within the data itself. The only thing that we know for certain is which customer tenant space it is coming from. For these reasons, tyGraph has taken the position of educating our customers on how tyGraph works so that they can advise us on their compliance needs. tyGraph offers several deployment scenarios to ensure our customers stay in compliance with local legislation and that cyber security vulnerabilities are minimized.

Please note if you are using a PDF, please download the latest version here:
https://datadictionaryimages.blob.core.windows.net/tygraphpublicsupport/tyGraph%20Compliance%20White%20Paper.pdf

About the Company

tyGraph is an award-winning Independent Software Vendor and makers of the flagship Office 365 Analytics solution by the same name. The company was formed in 2007 with the vision to help businesses drive value through integrated business intelligence (BI) solutions. This vision has evolved and now includes the use of Machine Learning (ML) and Artificial Intelligence (AI) platforms to extract the strongest signals from the billions of rows of data processed for our customers each month. Many Fortune 500 companies around the globe trust tyGraph to deliver the analytics they need so they can take actionable insights on the data presented. tyGraph is headquartered at 22 King Street, Suite 300, Waterloo, Ontario, Canada, N2J 1N8.

The services we provide are continually evolving, and so is the security threat landscape. To stay current, the company regularly discusses security matters in daily meetings with all development and operations staff.

Data Processing Centers

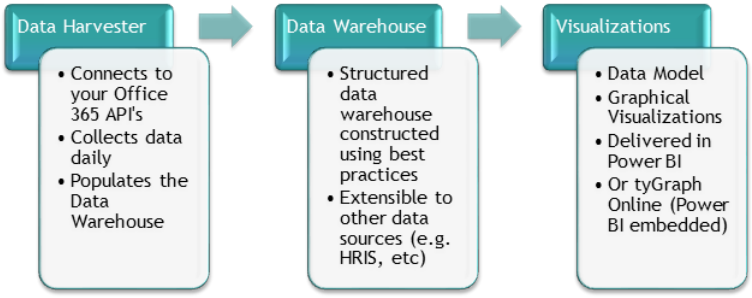
tyGraph has no physical server footprint outside of the Microsoft Azure data centers. We have done this to minimize our attack surface. To further minimize security risk, tyGraph leverages Microsoft Azure PaaS services (SQL Azure and Azure Web Jobs). This means that there are no underpinning servers subject to a patching cadence, exposed to virus vulnerabilities, or exposed to potential attack. These servers are all managed by Microsoft in Azure. Our default Azure hosting region is Azure North America, but we can host in other Azure regions such as the EU for GDPR compliance. These are available at the customer's direction.

Azure also affords tyGraph with key audit standards like ISO 27001, SOC (I, II, and III), and HIPAA to name a few. The full spectrum of compliance standards are documented here: **https://www.microsoft.com/en-us/TrustCenter/Compliance/default.aspx**.

tyGraph employee access to the Azure environment is controlled with named IP addresses and 2-factor authentication on AAD accounts. Access is only provided to those that need it for customer work.

How tyGraph Works

tyGraph consists of 3 foundational components:



Data Isolation and Securitization

Each component is dedicated to each customer in the full SaaS version of tyGraph. This design feature means that data is never co-mingled, as it is with many multi-tenant applications. tyGraph leverages Azure Platform-as-a-Service (PaaS) services, namely Azure App Services for the harvest components and SQL Azure for the database components. Using the PaaS environments ensures that the underlying servers are not exposed, minimizing the attack surface. The other advantage is that this environment is not prone to the injection of viruses. The security and compliance of this environment is under the

control of Microsoft. Given that our customers have already approved Azure as part of their existing compliance scenarios, it is much easier to approve an application like tyGraph that is domiciled in the same environment.

Azure Advanced Data Security Services

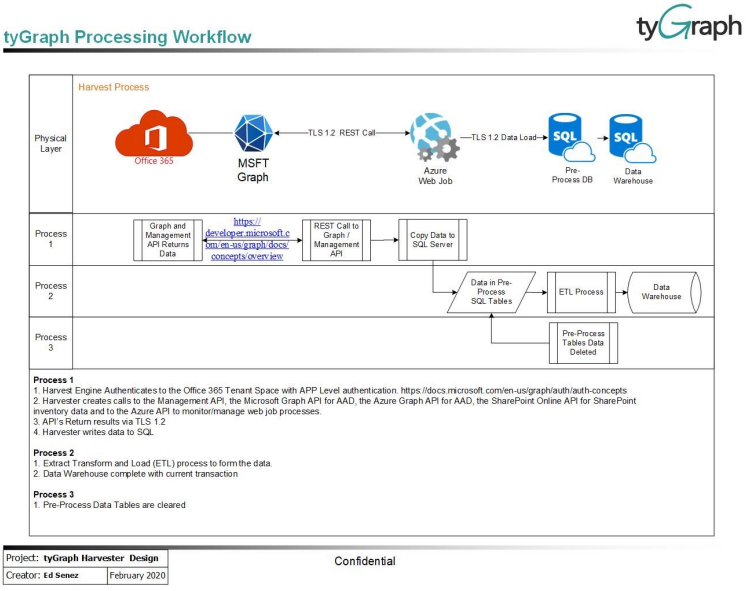
tyGraph subscribes to the additional Advanced Data Security offered by Azure. The following services are employed:

- All databases are monitored at a 'server' level: (i) Data Discovery & Classification; (ii) Vulnerability Assessment; (iii) Advanced Threat Protection.
- All Advanced Threat Protection settings are enabled (SQL injection; SQL injection vulnerability; Data exfiltration; Unsafe action; Brute force; Anomalous client login).
- Auditing is enabled and logs are stored in 2 locations.
- Risk assessment reports are delivered weekly.

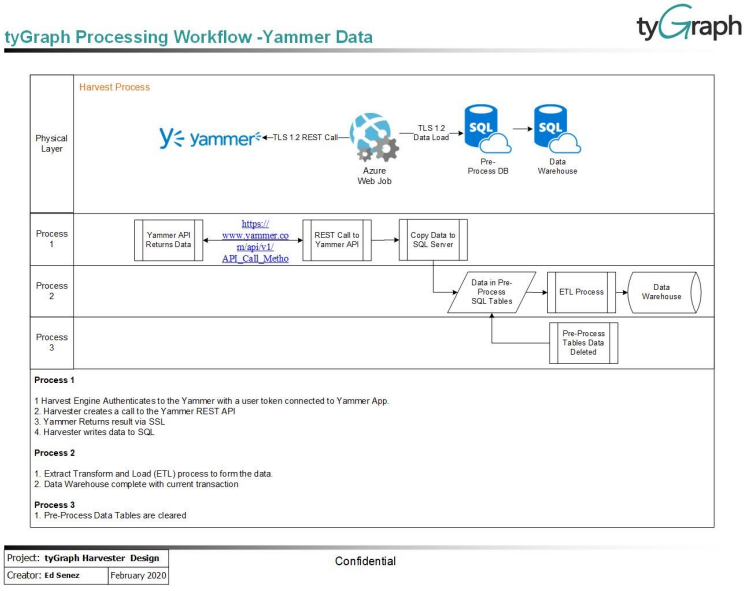
Data Transportation

The most common security question is the encryption status of the data throughout the lifecycle of the data harvest process. In short, data is encrypted throughout the lifecycle, including at rest. Microsoft enforces certain security protocols for the REST API's like TLS 1.2 for the transportation layer.

The following diagram depicts the data lifecycle for processing Office 365 data:



Yammer has a slightly different deployment because the data is not in the Microsoft Graph. Rather, Yammer (currently) uses a legacy "Export API" model. Some customers augment their Yammer data with Azure Active Directory (AAD) data. This data collection would follow the same process as the one noted above using the Graph API's.



Microsoft APIs harvested by tyGraph product family

We use the following Microsoft APIs as part of tyGraph:

- Microsoft Graph API
- Office 365 Management API
- SharePoint REST API
- Yammer REST API
- Yammer Data Export API
- Microsoft Text Analytics API

The above is the comprehensive list for customers using tyGraph Enterprise with all features turned on. For customers using select products, we would only use the API's that apply. Customers can choose to opt-out of some functionality, such as Text Analytics. We can also scope AAD data collection to certain specified fields and the direction of the customer.

Here is a detailed list of the APIs used by each product family and the specific API endpoint called for each:

tyGraph Pulse

SharePoint Online web

- site collections,
- online root webs,
- sites

Microsoft Graph Usage Reports

- graph all usage reports

Microsoft Graph

- graph users,
- graph user licenses,
- graph subscribed SKUs,
- graph groups,
- graph group owners,
- graph group users,
- graph user drives,
- graph group drives,
- graph organization domains

Yammer Export API / Several Yammer endpoints

- yammer unified groups (optional)

tyGraph for Teams

SharePoint Online web

- site collections,
- sponline root webs,
- sites

O365 Management

- o365 mgmt teams activities

Microsoft Graph Usage Reports

- graph usage report teams device usage,
- graph all usage reports,
- graph usage report teams user activity

Microsoft Graph

- graph users,
- graph user licenses,
- graph subscribed SKUs,
- graph groups,
- graph group owners,
- graph group users,
- graph user drives,
- graph group drives,
- graph organization domains,
- graph teams,
- graph team channels,
- graph team channel tabs,
- graph team channel members,
- graph team channel drives,
- graph team chat messages,
- graph team chat replies,
- graph teams message sentiments
- graph call record (optional)

Yammer Export API / Several Yammer endpoints

- yammer unified groups (optional)

tyGraph for OneDrive

O365 Management

- o365 mgmt onedrive activities

Microsoft Graph Usage Reports

- graph usage report onedrive usage

Microsoft Graph

- graph users,
- graph user licenses,
- graph subscribed SKUs,
- graph user drives

Yammer Export API / Several Yammer endpoints

- yammer unified groups (optional)

tyGraph for SharePoint

SharePoint Online web

- site collections,
- sponline webs,
- sponline sub webs,
- sponline lists,
- sponline wiki page library list items

O365 Management

- o365 mgmt AAD activities,
- o365 sharepoint activities,
- o365 general activities

Microsoft Graph Usage Reports

- graph usage report sharepoint activity,
- graph all usage reports sharepoint site usage

Yammer Export API / Several Yammer endpoints

- yammer unified groups (optional)

Microsoft Graph

- graph users,
- graph user licenses,
- graph subscribed SKUs,
- graph groups,
- graph group owners,

- graph group users,
- graph user drives,
- graph group drives,
- graph organization domains

tyGraph for Yammer

O365 Management

- o365 mgmt yammer activities

Microsoft Graph Usage Reports

- graph yammer usage reports

Yammer Export API / Several Yammer endpoints

- Yammer networks,
- Yammer users,
- Yammer groups,
- yammer unified groups,
- yammer group members,

- yammer activities,
- yammer export,
- yammer threads,
- yammer thread activities,
- yammer files,

- yammer likes,
- yammer topics,
- yammer message sentiments

Microsoft Graph

- graph users,

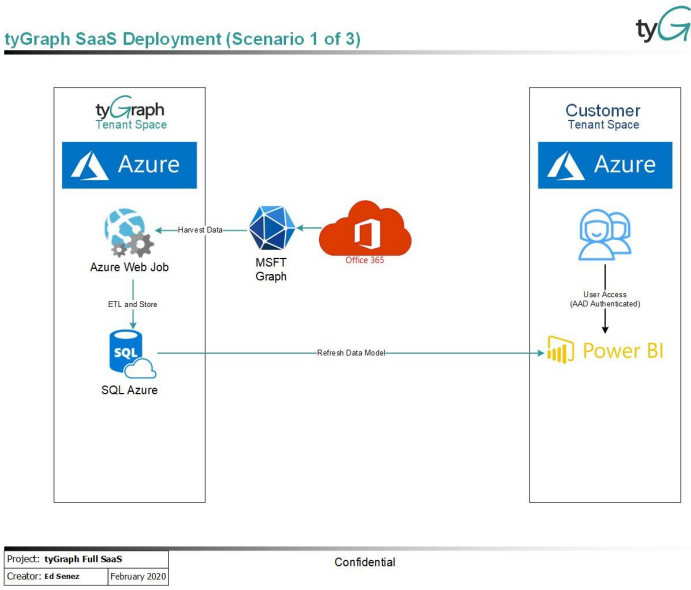
- graph user licenses,
- graph subscribed SKUs

Deployment Scenarios

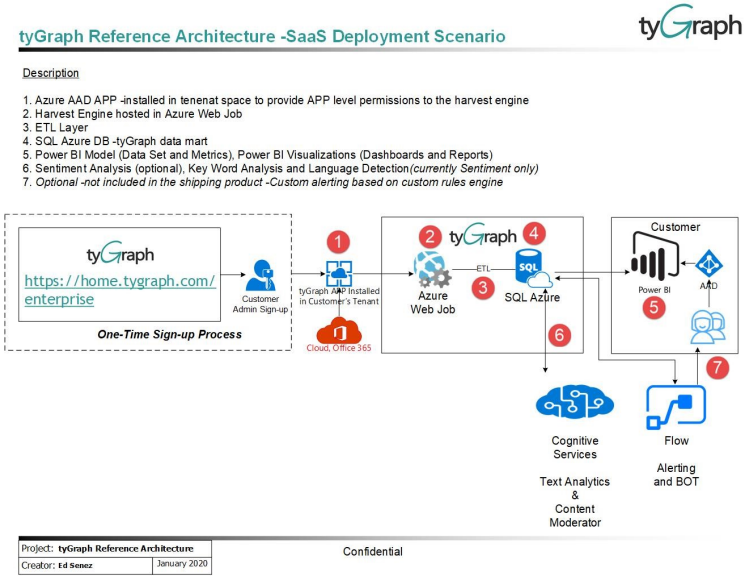
We offer customers three deployment scenarios with a few options within each scenario, as detailed below. We price all deployment scenarios the same way. If the customer is using their own tenant space, the cost of Azure services is paid for by the customer on their subscription.

Scenario 1 -tyGraph SaaS Deployment

Most customers use the tyGraph SaaS model wherein the entire production process lives within the tyGraph Azure Data Center. The default Azure Data Center for tyGraph is in the USA; however, we can choose any Azure data center at the customer's direction. Typically, EU-based customers have their data domiciled in the EU.
Again, each customer is assigned a dedicated Azure Web Job and a dedicated SQL Azure Database wherein all data is encrypted throughout the process, including at rest. These are PaaS services and not VM's.

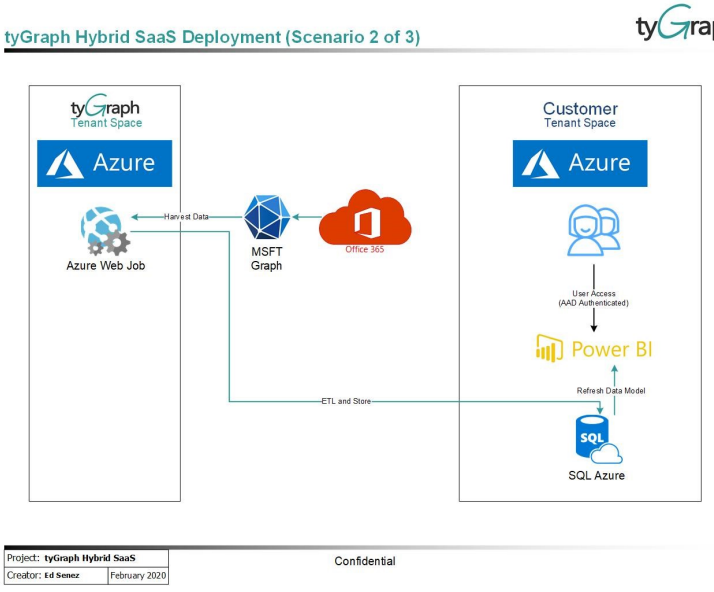


Reference Architecture for SaaS



Scenario 2 -tyGraph Hybrid Deployment

In this use case, the SQL Azure Database is in the customer's Azure tenant space while processing is completed within the tyGraph Data Center. Given that processing data is not persisted in the tyGraph data center, this scenario complies with most data residency legislation or internal governance requirements.

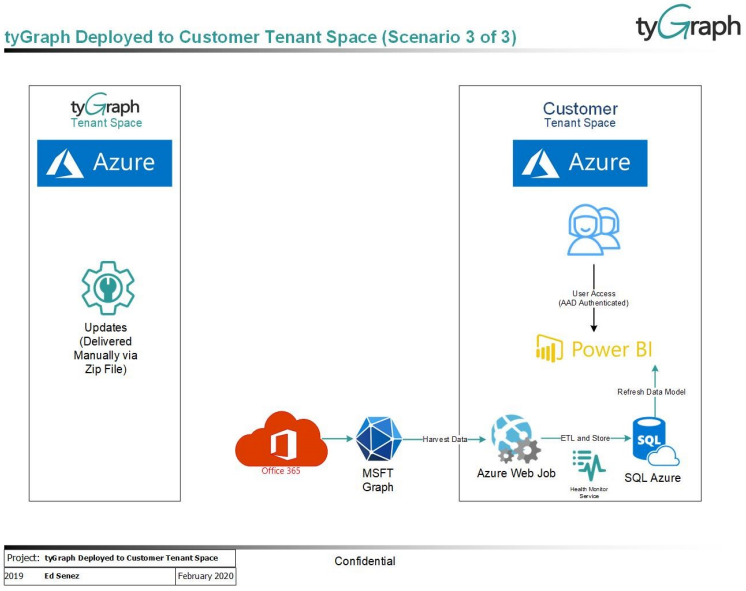


Scenario 3 -tyGraph Customer Deployed

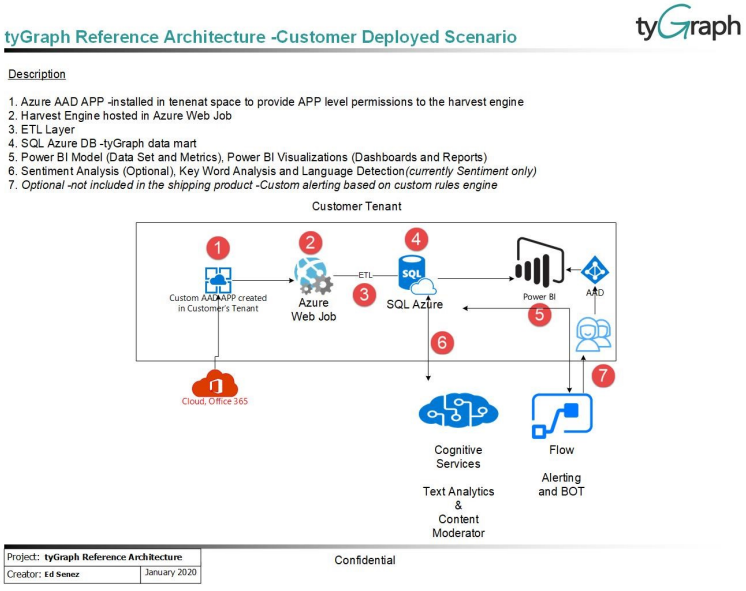
The preference for this scenario is to deploy using PaaS services in the customer's Azure tenant space (Azure Web Job and SQL Azure). However, we do have provisions to run on Virtual Machines. The Web Jobs are replaced with Windows Services on an IIS Server for the VM scenario. We have also deployed to On-Premises servers in special cases, but it is not a preferred alternative.

This deployment model provides for maximum security control and data residency. However, it does create some minor servicing challenges that are worked through with each customer on a case-by-case basis.

Please contact us for a copy of our tyGraph Installation Run Book for installation details.



Reference Architecture for Customer Deployed



The Sign-up Experience and AAD App Level Permissions

The Microsoft Graph requires application-level permissions to authorize access. Details can be found here: **https://docs.microsoft.com/en-us/graph/permissions-reference**

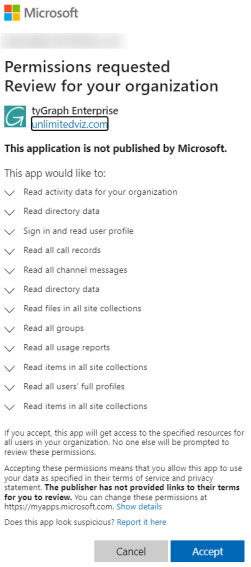
Each tyGraph product has its own set of permissions associated with it. See the **Permission Break-down** section of this article.

The path to signing up depends on the hosting solution chosen. Work with your tyGraph sales contact, tyGraph CSM, or partner for instructions on how to sign-up. Should you choose tyGraph-hosted, you will need to authorize permission to call the APIs. **Only a tenant administrator can complete the authorization steps.** It should also be noted that the user account used to authorize the App is not tied to tyGraph in any way. Typically, you will be prompted to log in with your Office 365 Admin account after selecting the link.

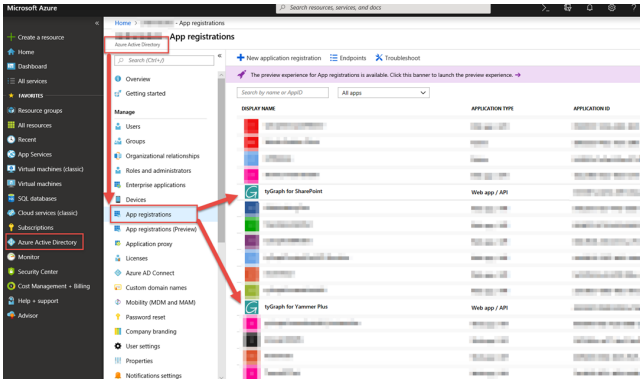
For example, an Enterprise authorization will bring you to the following Microsoft Authorization Page. This screen capture is for **tyGraph Enterprise**, which includes all API points used by the tyGraph App (except Yammer because it has a different API access authorization).

Each tyGraph product will use a different set of API endpoints based on the scope of the Microsoft workload in scope (i.e., Teams, SharePoint, OneDrive, etc).

Important: **If you chose Enterprise, you do not need to authorize the other tyGraph Apps.**



Once installed, these Apps will be registered in your AAD like the screenshot below.



For customers who purchase tyGraph Enterprise, only 1 App is required.

Customers who opt for Scenario 3 (Customer Deployed), cannot use the tyGraph Apps. As part of the installation process, a Custom App will be created. Details can be found in our tyGraph Installation Run Book (available upon request).

Permission Break-down

Below are all the permissions required, where they are used, and why they are required. Descriptions of each can be found here: **https://docs.microsoft.com/en-us/graph/permissions-reference**

Azure Active Directory Graph API

Permission: **Directory.Read.All**
Display string: *Read directory data*
Primary features: The purpose of collecting data against AAD is to help provide HRIS details in the reports. Here are the fields we ingest; we can collect all or none, at the Customer's discretion:
Email, Address, City, ZipPostalCode, StateProvince, Country, CountryCode, Continent, Building, Organization, Department, JobTitle, Division, BusinessUnit, ManagerEmail, Latitude, Longitude, OrgLevel01 - OrgLevel10

Microsoft Graph API

Permission: **ChannelMessage.Read.All**
Display string: *Read all channel messages*
Primary features: Channel messages is the underpinning of the Teams reporting. It is how we determine Channel activity. If this permission is not granted, most of the Teams reporting will not work. We would only be able to provide what is in our Pulse product for Teams.

Permission: **Directory.Read.All**
Display string: Read directory data
Primary features: The Directory data provides the ability to mash-up HRIS data with the Microsoft 365 data. If this permission is not granted, you would not have 365 insight.

Permission: **Files.Read.All**
Display string: *Read files in all site collections*
Primary features: Read Files in the Site Collection is necessary for us to identify files. Microsoft makes this permission overreaching because it allows us to access the entire file, but this is something we do not do. We do this to inventory files and then we match inventory with the activity against these files. This is fundamental to the tyGraph for SharePoint product.

Unfortunately, Microsoft does not give us a more granular capability of scoping to certain data. Even though we do not collect files, the capability is there, which is a concern and something we are pushing Microsoft to update. **tyGraph for SharePoint - Product Guide - 2020 Release : Support**

Permission: **Group.Read.All**
Display string: *Read all groups*
Primary features: We need to inventory the groups for reporting purposes. Without this, there will be a lot of unmatched information.

Permission: **People.Read.All**
Display string: *Read activity data for your organization*
Primary features:

Permission: **Reports.Read.All**
Display string: *Read all usage reports*
Primary features: **Pulse Product guide**

Permission: **Sites.Read.All**
Display string: *Read items in all site collections*
Primary features: Similar to Read Files, but to list items and not files. This is to get all the sites that are in the tenant. **Sharepoint Product Guide**

Permission: **User.Read**
Display string: *Sign in and read user profile*
Primary features: These are required as core for all of our products. We use certain details to fill in where there would otherwise be “blanks” or data summarized to “blank”.

Permission: **User.Read.All**
Display string: *Read all users’ full profiles*
Primary features: These are required as core for all of our products. **User Attribute Options**, **How users work in tyGraph**

Permission: **CallRecords.Read.All**
Display string: *Read all call records*
Primary features: Single peer-to-peer call or a group call between multiple participants, sometimes referred to as an online meeting.

Office 365 Management API

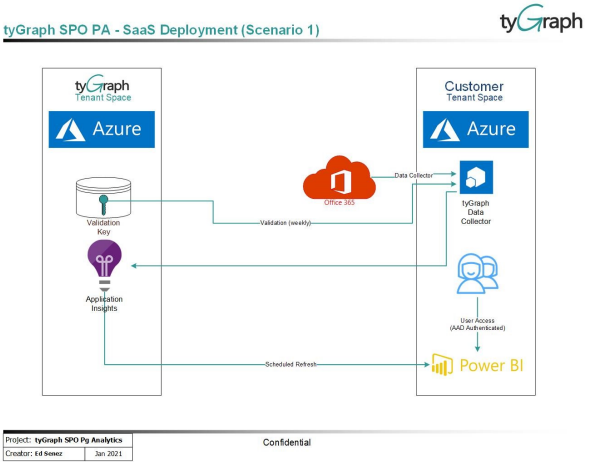
Permission: **ActivityFeed.Read**
Display string: *Read activity data for your organization*
Primary features: An aggregation of actions and events for specified content types as Azure Active Directory, SharePoint, OneDrive, Teams, or Yammer. **tyGraph for SharePoint - Product Guide - 2020 Release : Support**

tyGraph Pages

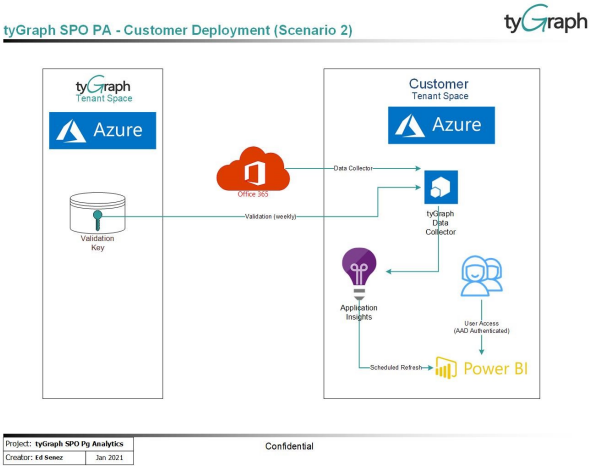
With tyGraph Pages, we collect data different than the processes above. A separate chapter is provided to articulate the architecture and deployment scenarios. The major differences between the architectures are as follows:

1. Data is stored in Application Insights instead of SQL Server
    a. For larger or highly active sites (>500M transactions), Application Insights is replaced with Azure Data Explorer (ADX).
2. Obfuscation is enforced at the point of data collection.
    a. PII data such as employee name and email are either included in collection or blocked from collection by the tyGraph Pages Engine
    b. The customer has three obfuscation choices.
        i. Clear - no obfuscation
        ii. Fully obfuscated
        iii. Partial obfuscation based on user’s UsageLocation country code assigned with his or her Office 365 license.
3. Instead of collecting data from an API, the tyGraph Pages Engine is deployed as an SPFx solution into the SharePoint tenant and scoped either to the whole tenant or a portion of it.
4. Data is refreshed in near real-time (approximately a 5-minute delay).
5. Customers can optionally grant the User.Read Graph permission to allow filters based upon user Azure Active Directory attributes in reports

Like the other tyGraph products, the customer can host the data with tyGraph or in their own tenant space. Scenario 1 depicts our SaaS solution where the data is hosted at tyGraph.



Scenario 2 depicts a scenario where tyGraph Pages is hosted entirely in the customer’s tenant space.



Data Ownership

Customers own all their data always. If the service is terminated, the customer can request that the data be provided to them in an SQL database. This data will be provided at no additional charge. Once this is complete, the data will be deleted according to our Data Destruction Policy, and a certificate of deletion will be issued.
Customers may choose to participate in the tyGraph benchmarking program. In this case, select data is anonymized, aggregated together and provided back to the customer in a dashboard through tyGraph Online. Only customers who chose to participate in the benchmarking offering will be given access to this dashboard. This is a voluntary Opt-In program and the default setting for customers is Opt-Out.

Data Residency

Data can be stored in either tyGraph's default Azure data center in the USA, or a data center in other geographies such as the EU for GDPR purposes. If the customer is hosting their own data, then it is assumed that they will geo-locate the data appropriately.

GDPR Compliance

Overview

tyGraph offers many capabilities and options to either opt out users from any data being collected or by obfuscating their personally identifiable information (PII). tyGraph can obfuscate PII information such as the user's name, email address and photo from appearing in the visualizations. The PII information is either entirely removed, as in the case with the photo or overwritten with random numbers as in the case with use names.

Applicable Products

tyGraph Data Protection settings are applied uniformly across all compatible licensed tyGraph products for your organization.

Data Protection supported products:

- tyGraph for Teams
- tyGraph for SharePoint
- tyGraph for Yammer
- tyGraph Pulse

Currently Unsupported:

- tyGraph for One Drive
    - Note: customers often use this product for Data Loss Prevention (DLP) such as tracking external file shares and obfuscating the user defeats the purpose of the product. For this reason, we have not included data protection settings in tyGraph for OneDrive.

Data Protection Levels

There are three levels of data protection. Each level is mutually exclusive.

Un-obstructed: No logic is applied to hide or remove user's information or activity. A user's attributes can be viewed in reports and their data is counted exactly as in reality.

Obfuscation: All data is collected and counted. All identifying information about that user is replaced with a random number.

Opt-Out: Only data generated by opted-in users is reported. Personal and activity data for opted-out users is entirely removed from reporting. This removal is far reaching and effects the products in the following ways:

- Transactions:
    - Teams, SharePoint, and/or Yammer activity where the user is opted out is not reported.
- tyGraph for Yammer:
    - Examples:
        - Messages and likes on messages that the user posted are removed.
        - Likes made by non-opt-out users on a message posted by an opted-out user are removed.
          Example, William (not opted out) likes Martha's (opted out) message. William's like is not counted for William.
        - Praises given to a non-opt-out user by an opted-out user are not counted and vice versa.
    - Complete List:
        - Users
        - Activities
        - Files
        - Group Members
        - Likes
        - MessageAttachments
        - MessageXref

        - Messages
        - MessageShares
        - Pages
        - Praises
        - ThreadTopics
        - tyGraphURLDetail
        - UserMentions
        - UserPraises

- tyGraph Pulse:
    - Objects owned by an opted-out user are removed.
    - Groups created opted out user are removed
- tyGraph for SharePoint:
    - Any activity taken by an opted-out user is removed.
    - Any objects created by an opted-out user is removed.
    - Any SP webs where the owner is opted out are removed.
    - If Group Owner is opted out, data is not included.
    - If Site Owner is opted out, data is not included.
    - Activity Inventory URLs on users' objects

So long as one user on a file is not opted out the object will be reported. If all users on an object opt out, the file will be removed.

The Right to be forgotten

The largest applicability of this GDPR compliance requirement is specific to tyGraph for Yammer or tyGraph for Teams where individuals are part of the network and users have posted messages into the network. Under GDPR legislations, individuals have the 'right to be forgotten'. Recall that tyGraph is a downstream system, so data purging is not valuable if the data is persisted upstream in Yammer or Teams. It would be re-collected on the next harvest cycle.
Hard deleting a user via the Yammer UI deletes all their messages in the Yammer backend, however the Yammer API endpoints do not automatically indicate that the messages have been deleted. Therefore, custom logic has been implemented to ensure all aspects of data is purged like what appears in the Yammer UI.

It is not possible to identify and delete messages where a user is mentioned in the content of a message. Most customers feel that this would be a far-reaching requirement that would result in the deletion of corporate knowledge and therefore is not in the scope of the GDPR legislation.

Conclusion

Security and compliance are an ever-evolving effort for our company. We continue to review and enhance our capabilities as part of our business operations. To that end, if there are requirements that you do not find available in tyGraph, please feel free to connect with us and we will work together to address your needs.