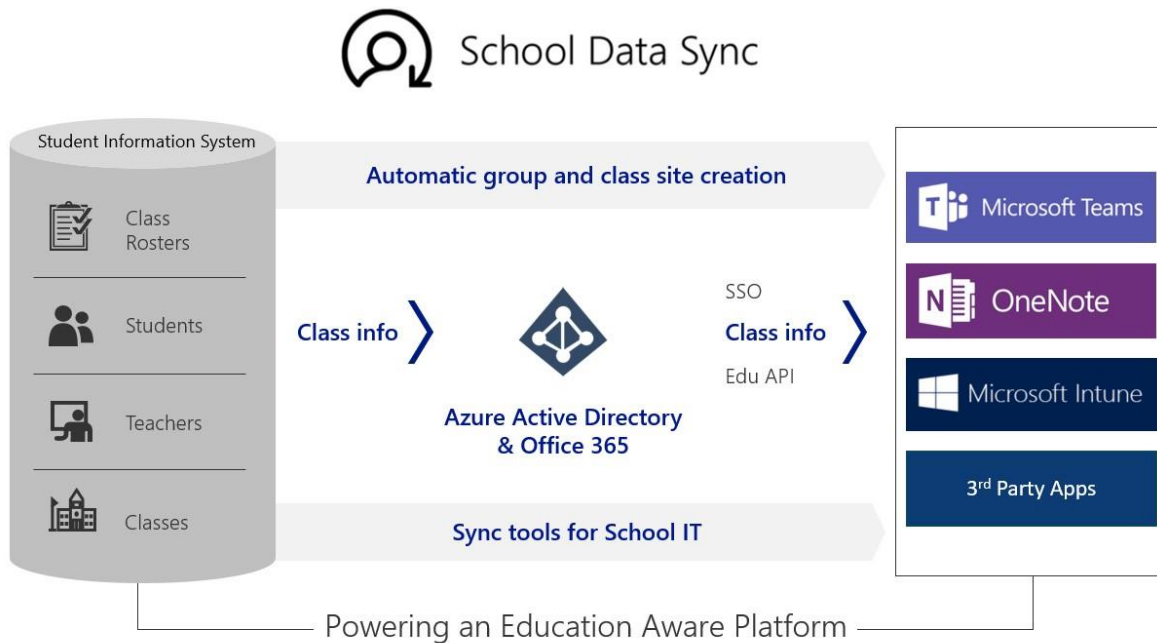


Options for Partners Integrating with School Data Sync Using Azure Data Factory

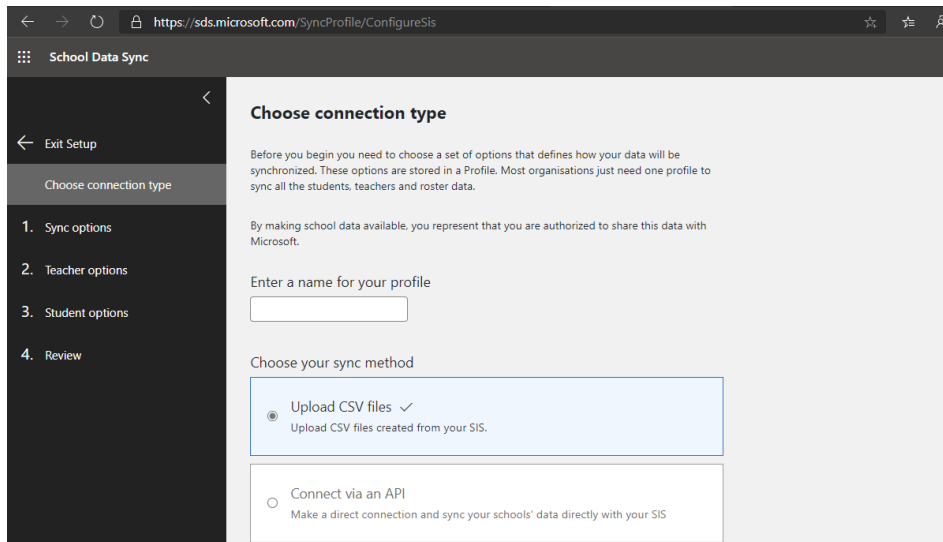
*© 2020 Microsoft Corporation. This document is for informational purposes only.
MICROSOFT MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS SUMMARY.
The names of actual companies and products mentioned herein may be the
trademarks of their respective owners*

Background

[Microsoft School Data Sync](#) is a free service in Office 365 for Education that reads the school and roster data from a school's Student Information System (SIS). It creates Office 365 Groups for Exchange Online and SharePoint Online, class teams for Microsoft Teams and OneNote Class notebooks, school groups for Intune for Education, and rostering and SSO integration for many other third-party applications.



SDS helps to automate the process of importing and synchronizing Student Information System (SIS) data with Office 365. In order to configure SDS, a school administrator can utilize the [SDS admin UI](#) within Office 365:



More information regarding integration options can be found here: [Deploying School Data Sync](#)

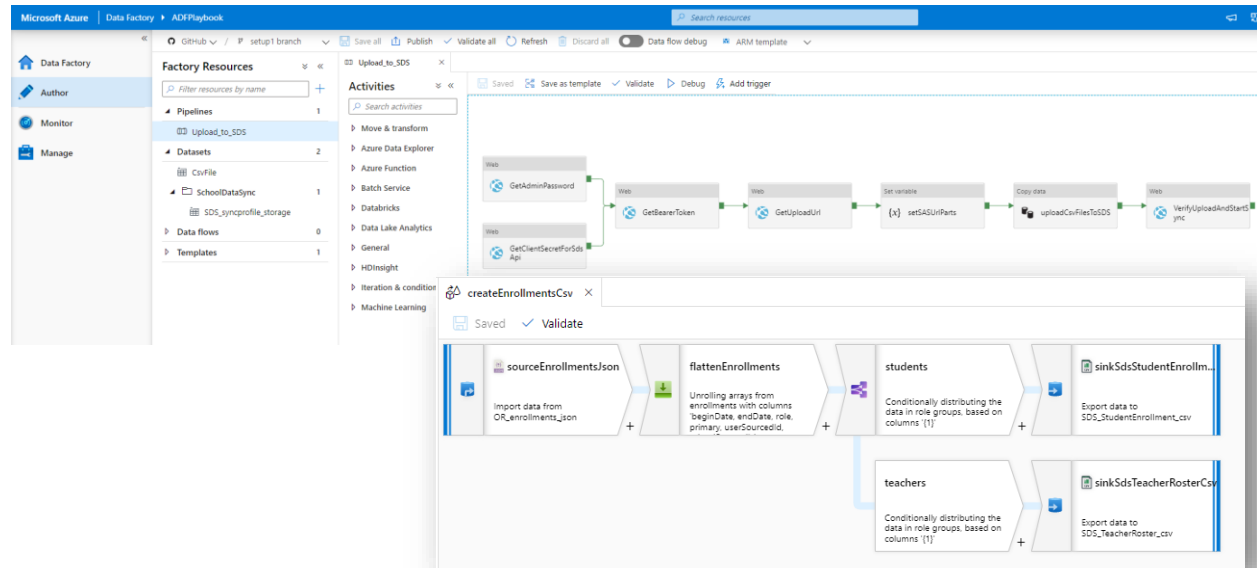
SDS also gives application developers an [SDS management API](#) that can be used within a third-party product to provide a tightly integrated SDS admin experience within that product. App developers can also use the [Education API](#) to access section, student, teacher, school, and class roster information to personalize experiences for teachers and students who use their apps.

More information on how to use the SDS management API can be found in the [SDSProfileManagementDocs github repo](#).

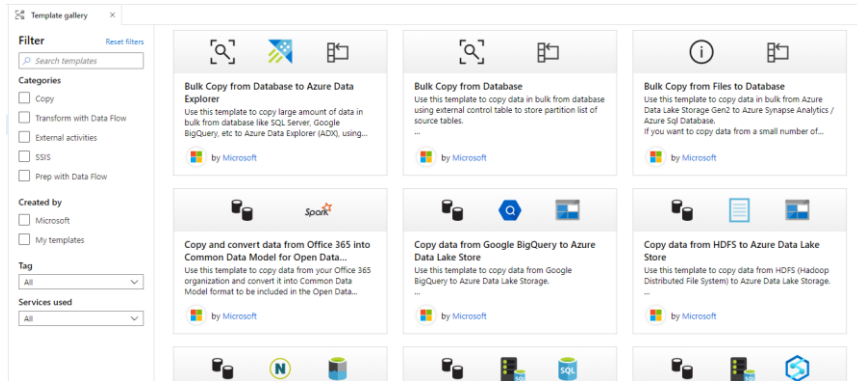
Another set of integration options leverages Azure Data Factory to provide a highly flexible, highly scalable, cost-effective, and secure approach that addresses complex and high-volume scenarios. This is an ideal option for customers or partners looking for a low-code/no-code integration approach that still provides the power and flexibility of leveraging the SDS management API.

Intro to Azure Data Factory

Azure Data Factory (ADF) provides an industrial strength data integration platform that is highly secure and fully managed in the cloud. ADF is a product that provides the ability to ingest, move, prepare, transform, and deliver your data in a few clicks, and complete your data integration projects within a web-accessible visual environment.



In addition to the large and growing feature set of ADF, there is an ever-growing catalog of data integration pipeline templates made available, allowing customers to leverage ETL pipelines that have been built according to best practices and fully tested.



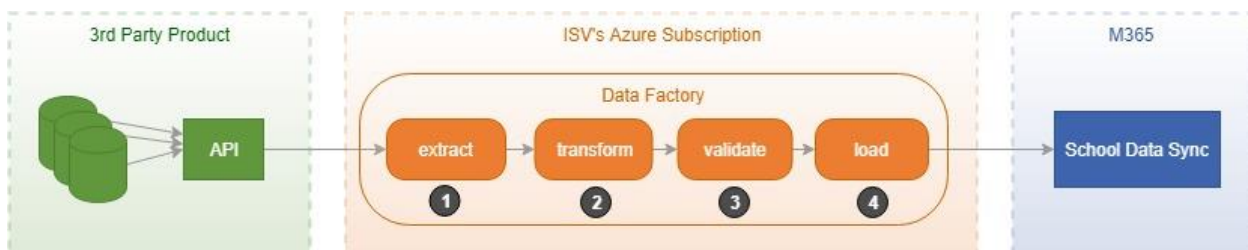
There are also a large number of system integrators who can provide custom solutions, either based on these pre-existing solutions or from scratch.

ADF Based Solutions for Integration with SDS

By combining the usability and deep feature set of ADF with the flexible SDS management API, a number of challenging integration scenarios can be effectively addressed in a way that is maintainable and easy to monitor. This section presents the high-level solution architecture for 3 scenarios.

Scenario #1: Automated sync via API

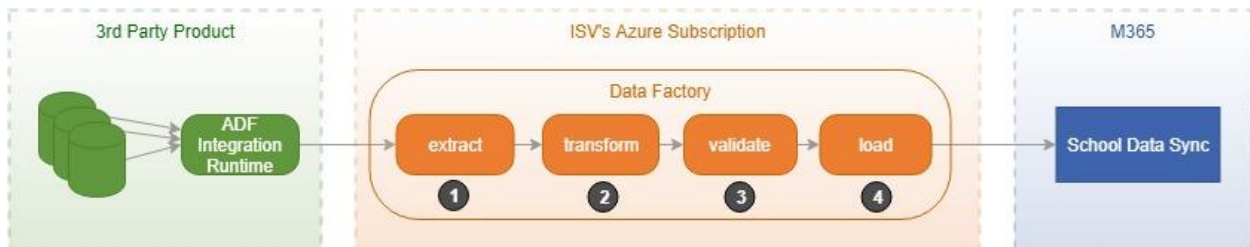
Partners who wish to offer an automated SDS integration directly from their product need to have a solution which can be scheduled to periodically perform an extract of the customer's data, perform transformations on the data to construct the SDS file format, then deliver the files to SDS via the management API. In the diagram below, an ISV has setup an integration pipeline within their ADF instance to facilitate the ETL process needed for each customer.



- 1) The integration pipeline in ADF invokes the product API to perform the data extract
- 2) The extracted data is transformed as needed, to create the data format needed by SDS
- 3) Validation checks are run, allowing for the customer to be alerted to any data issues
- 4) The data is delivered to SDS

Scenario #2: Automated sync via direct access

For partners who wish to perform an extract directly from on-premises resources or cloud resources within a VPN, ADF offers the [Integration Runtime](#).

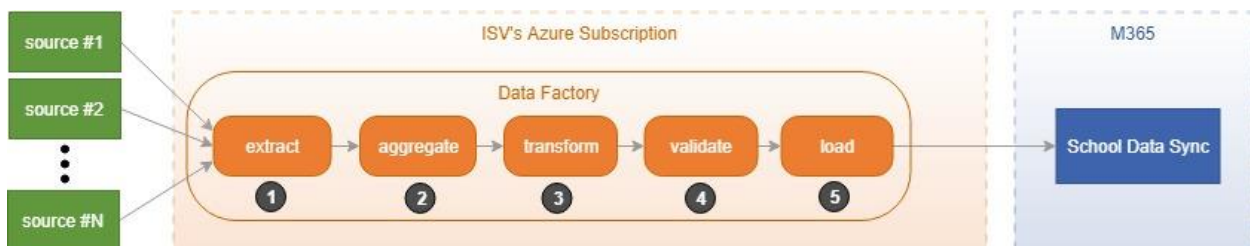


- 1) The integration pipeline in ADF securely connects to the Integration Runtime and performs the data extraction (this could be a direct db query, or picking up files from the file system, etc)
- 2) The extracted data is transformed as needed, to create the data format needed by SDS
- 3) Validation checks are run, allowing for the customer to be alerted to any data issues
- 4) The data is loaded to SDS

Note that any of the four stages depicted within ADF could be augmented or replaced by partner code as needed. For instance, a partner may prefer to have the ETL process occur within their product, and simply utilize ADF for stages 3 and 4 to facilitate the scheduled delivery of the extracts to SDS for each customer.

Scenario #3: Aggregation of multiple sources into a single tenant

SDS has a limit of 5 sync profiles per tenant. This means that for a school district that utilizes a single Office 365 tenant but has a large number of schools that each have their own SIS instance, there is a need to aggregate the data before delivery to SDS. For cases such as these, ADF can be used to pull in the extracts from each system and aggregate the data before then continuing on with the rest of the standard pipeline.



- 1) The extract stage cycles through and performs the extract on each source (the source systems may be on-premises or cloud or a combination)
- 2) The extracted data is aggregated (if necessary, id values are given a prefix based on the source id, to ensure uniqueness across source systems)
- 3) The extracted data is transformed as needed, to create the data format needed by SDS
- 4) Validation checks are run, allowing for the customers to be alerted to any data issues
- 5) The data is loaded to SDS

Note that if each school has a unique domain, in order to properly map extracted users to existing users in AAD, the new [SDS v2 CSV file format](#) must be used.