

ИСПОЛЬЗОВАНИЯ POWER FX ДЛЯ РАБОТЫ В COLLABORATION PLATFORMS

USING POWER FX FOR COLLABORATION PLATFORMS

*I. Voronkov
S. Saradgishvili*

Summary. This article is devoted to an example of organizing work on building ETL processes in an enterprise contour. This technique can serve as an illustration of the idea of two key trends in the development of IT technologies: the simultaneous growth of opportunities that are provided by certain software products in conjunction with a decrease in the entry threshold. The review is carried out on the Microsoft 365 platform. Separately, we note that this language is part of the Power Platform. This is a relatively new vision of Microsoft corporation about the business data warehouse as a single point of connection and storage of data, tools. Together with the Power platform authors researched the Power automate product — a workflow designer, partly the successor to SSIS, SharePoint workflow engine. These tools can automate enterprise activities. To enhance the ability to develop with third-party technologies and programming languages, Microsoft has introduced the Microsoft Graph API.

Keywords: ETL, Collaboration Platforms.

Воронков Илья Александрович

Аспирант, Санкт-Петербургский политехнический университет Петра Великого
iliaftk@outlook.com

Сараджишвили Сергей Эрикович

К.т.н., доцент, Санкт-Петербургский политехнический университет Петра Великого
SSaradg@yandex.ru

Аннотация. Данная статья посвящена примеру организации работы по выстраиванию ETL-процессов в контуре предприятия. Данная методика может служить иллюстрацией идеи о двух ключевых тенденциях в развитии ИТ-технологии: одновременный рост возможностей, которые предоставляют те или иные программные продукты в совокупности со снижением порога входа. Обзор выполнен на решении платформе Microsoft 365 как наиболее часто встречающегося решения по созданию Collaboration platform в учреждениях за рубежом. Отдельно отметим, что данный язык является частью Power platform. Это относительно новое видение корпорации Microsoft о бизнес-хранилище данных как о единой точке подключения и хранения данных, инструментов. Вместе с Power platform свое развитие демонстрирует продукт Power automate — конструктор рабочих процессов, отчасти наследником SSIS, SharePoint workflow engine. Вместе эти инструменты способны закрыть большинство задач по автоматизации деятельности предприятия. Для расширения возможности разработки с использованием сторонних технологий и языков программирования Microsoft представила Microsoft Graph API.

Ключевые слова: ETL, платформы для совместной работы.

Введение

ИТ рынок сейчас переживает новый виток своего бурного развития. Спрос на специалистов в этой области остается по крайней мере высоким [1]. Стоимость проектов и услуг по законам рынка растет вслед за спросом. Одним из вариантов решений задач для автоматизации и диджитализации стали многочисленные платформы, работающие по принципам low-code и zero-code [2]. Общий подход в таких решениях сводится к тому, что большинство процессов обработки могут быть представлены в виде графического конструктора, который работает по UML/BPMN. Основная особенность — разработчику не надо задумываться о том, как происходит взаимодействия на уровне структур данных, какие алгоритмы используются для выборки, сортировки, как распараллеливаются вычисления. Главная задача: реализация бизнес-требования, путем создания диаграммы состояния процессов.

С точки зрения бизнеса, данная парадигма — это наиболее эффективный способ решения задачи, так как бизнес-пользователи привыкли мыслить именно задачами, которые должна решать та или иная система. К сожалению, на данном этапе развития технологий мы не в состоянии полностью перейти на подобные технологические решения. Причины могут быть разными: невозможность полностью решить задачу, используя лишь средства конструктора, нехватка специалистов, обладающих навыками моделирования, программирования и бизнес-опытом, наличие большой прослойки legacy code. Тем не менее стоит отметить рост подобных решений не только в сфере бизнеса, но и в научной среде [3–4]. Одновременно с развитием технологий все острее становится вопрос о месте данных решений в будущем. Многие владельцы крупных ИТ-компаний продвигают идею того, что программирование и разработка софта должно перестать быть узкоспециализированной областью [5]. Основная мысль сводится к тому,

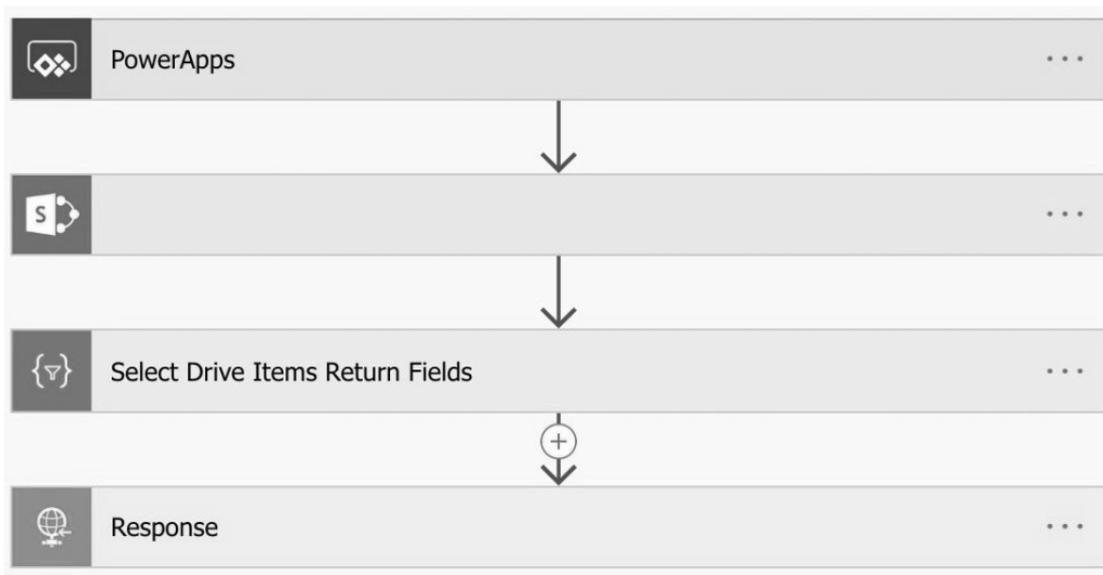


Рис. 1. Общая схема Flow

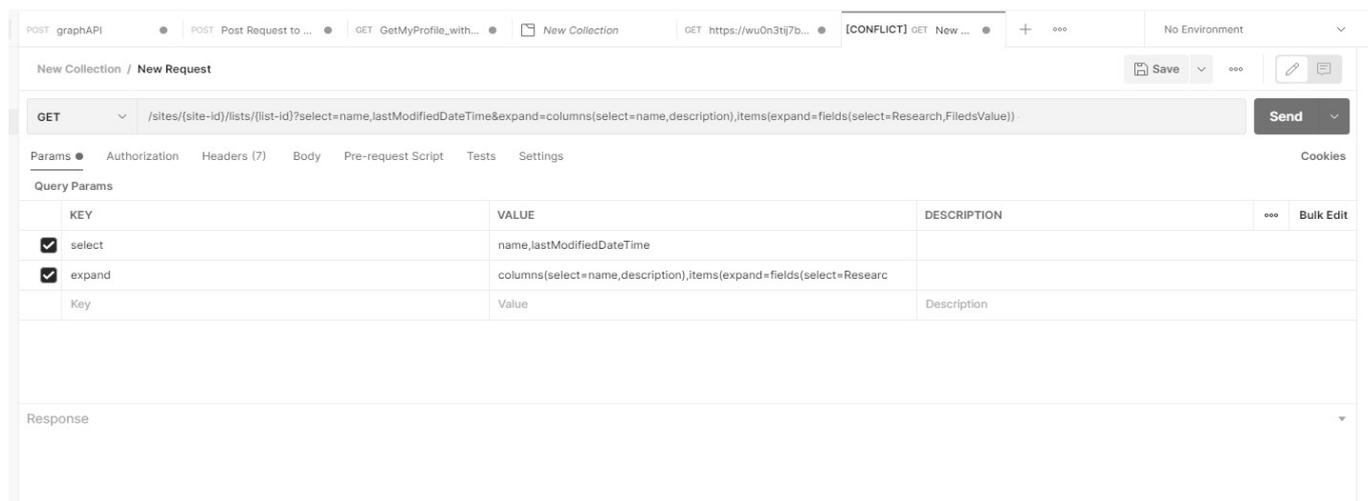


Рис. 2. Схема GET запроса

что в будущем люди будут решать с помощью программных инструментов свои персонализированные задачи. Для этого вам не придется получать образование, проходить специальные курсы. Все, что вам следует знать, это то, чего вы хотите добиться с помощью программы, и как она может облегчить вам жизнь. Конечно, подобные идеи могут казаться утопическими и несбыточными на горизонте ближайших 10–20 лет, но, оглядываясь на историю развития ИТ-отрасли мы начинаем думать, что все это может произойти уже в ближайшем будущем [6]. Рассмотрим задачу по настройке ETL процесса из корпоративного хранилища данных. Существует несколько подходов по организации работы подобного

рода[7]. Большинство из них сводятся к серьезным вложениям средств на этапе проектирования, разработки, приемке. В данной работе используется механизм Low-code работы, который становится все более популярным за счет низкой стоимости создания, но высокой операционной стоимости эксплуатации[8][9]. В обзоре представлен механизм получения элементов из произвольного хранилища, роль которого выполняет SharePoint. Адрес хранилища (сайт коллекция) может динамически изменяться, получая новый URL в качестве входного параметра. Задача состоит в универсальном подходе получения как данных из источника (бинарные файлы), так и извлечению информации метаданных.

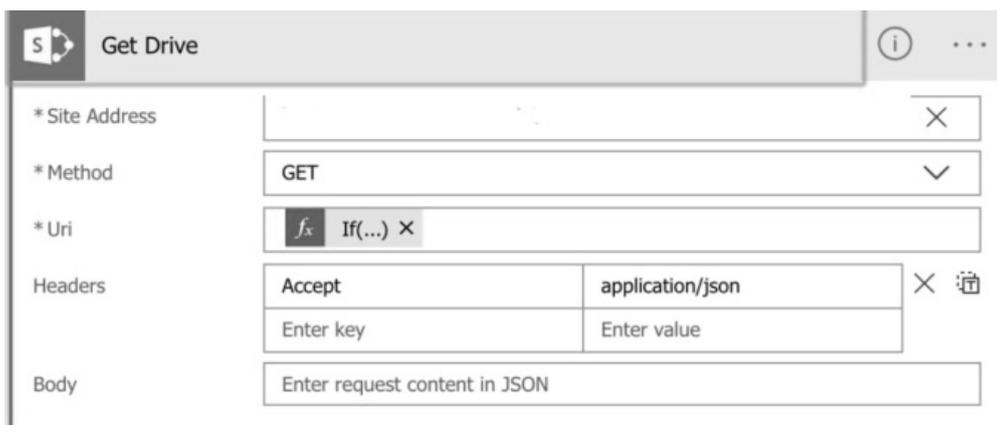


Рис. 3. Динамическое подключение к SharePoint

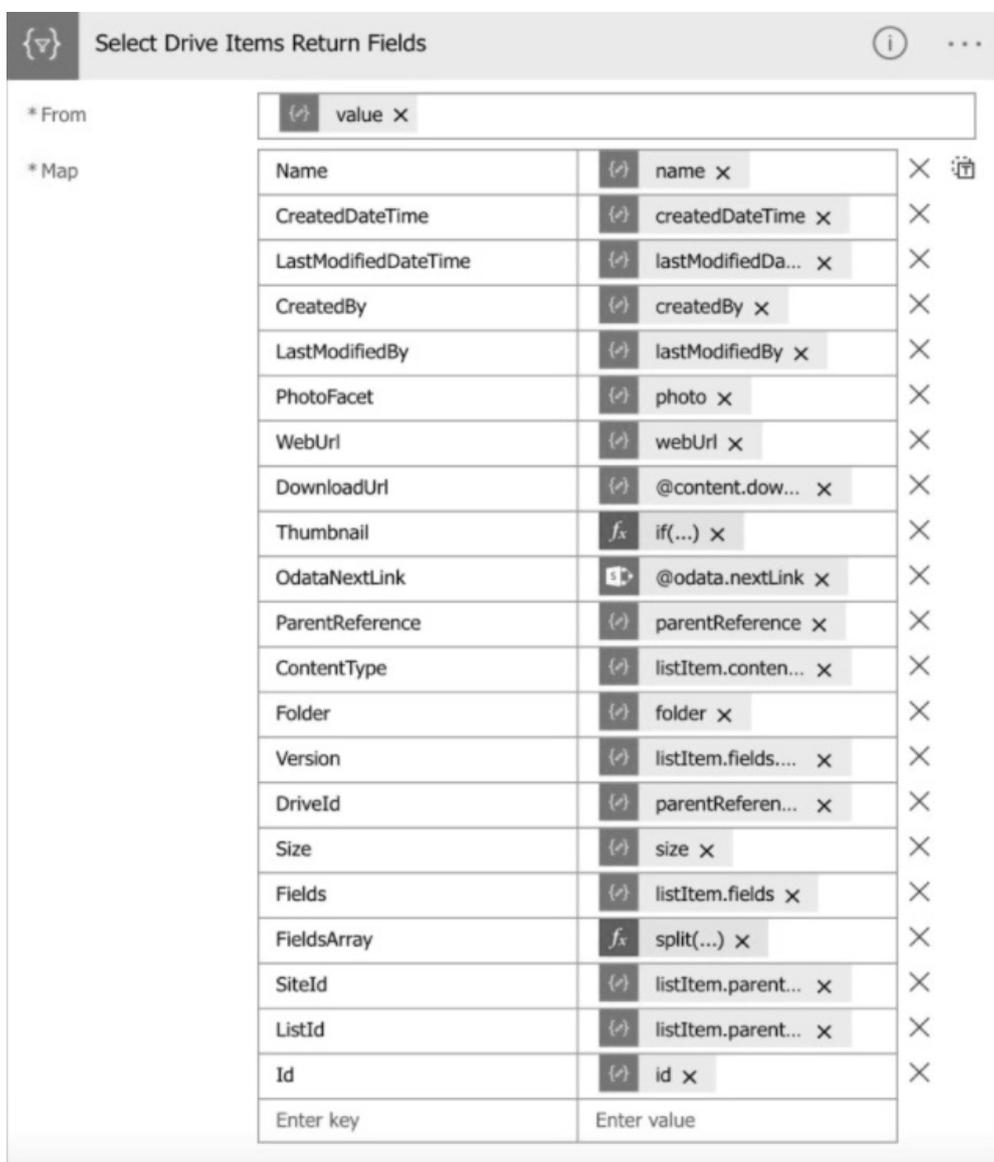


Рис. 4. Динамическое подключение к SharePoint

Таблица 1. Описание полей

Свойство	Функция	Описание
OdataNextLink	body('Get_Drive_Items')?['@odata.nextLink']	Добавление автоматической пагинации страниц
DownloadUrl	item()?['@content.downloadUrl']	Это свойство предоставляет (временный) URL-адрес с предварительной аутентификацией для каждого файла, возвращаемого в наборе результатов выполнения Flow.
Thumbnail	if(empty(item()['thumbnails']),, item()['thumbnails'][0]['large']['url'])	Получение свойства предпросмотра документа
FieldsArray	FieldsArray	Динамический набор полей метаданных

Learn about working with collections

AbsoluteUri	DisplayName	Id	Value
	MIG_280x60-1.png	%252fLists%252fBP_nominee%252fAttachments%252f75%252fMIG_280x60-1.png	appres://datasources/BP_nominee/table/179ebeeec-165e-471d-92b2-ec22b9fa4259/rows/75/reference/https:%2F%2Fеurope-002.azure-apim.net%2Fapim%2Fsharepointonline%2F77e32e0e6f0c1c10015e1f17f150
https://[redacted].sharepoint.com/sites/PowerAppsSandbox/Lists/BP_nominee/Attachments/75/sample.pdf	sample.pdf	%252fLists%252fBP_nominee%252fAttachments%252f75%252fsample.pdf	appres://datasources/BP_nominee/table/179ebeeec-165e-471d-92b2-ec22b9fa4259/rows/75/reference/https:%2F%2Fеurope-002.azure-apim.net%2Fapim%2Fsharepointonline%2F77e32e0e6f0c1c10015e1f17f150
https://[redacted].sharepoint.com/sites/PowerAppsSandbox/Lists/BP_nominee/Attachments/75/script.txt	script.txt	%252fLists%252fBP_nominee%252fAttachments%252f75%252fscript.txt	appres://datasources/BP_nominee/table/179ebeeec-165e-471d-92b2-ec22b9fa4259/rows/75/reference/https:%2F%2Fеurope-002.azure-apim.net%2Fapim%2Fsharepointonline%2F77e32e0e6f0c1c10015e1f17f150
https://[redacted].sharepoint.com/sites/PowerAppsSandbox/Lists/BP_nominee/Attachments/75/Voronkov_Saeadzhashvilli.docx	Voronkov_Saeadzhashvilli.docx	%252fLists%252fBP_nominee%252fAttachments%252f75%252fVoronkov_Saeadzhashvilli.docx	appres://datasources/BP_nominee/table/179ebeeec-165e-471d-92b2-ec22b9fa4259/rows/75/reference/https:%2F%2Fеurope-002.azure-apim.net%2Fapim%2Fsharepointonline%2F77e32e0e6f0c1c10015e1f17f150

Рис. 5. Коллекция значений исполненного потока

```

{
  "inputs": {
    "host": {
      "connection": {
        "name": "@json(decodeBase64(triggerOutputs().headers['X-MS-APIM-
Tokens']))['$connections']['shared_teams']['connectionId']"
      }
    },
    "method": "post",
    "body": {
      "recipient": "@triggerBody()['Postmessageinachatorchannel_Recipient']",
      "messageBody": "<p>@{triggerBody()['Postmessageinachatorchannel_Message_1
']} shared with you<br>\n<br>\n<a href=@{triggerBody()['Postmessageinachatorchannel_M
essage']}>@{triggerBody()['Postmessageinachatorchannel_Message_2']}</a>\n</p>"
    },
    "path": "/beta/teams/conversation/message/poster/Flow bot/location/@{encodeUR
IComponent('Chat with Flow bot')}",
    "authentication": {
      "type": "Raw",
      "value": "@json(decodeBase64(triggerOutputs().headers['X-MS-APIM-
Tokens']))['$ConnectionKey']"
    }
  },
  "metadata": {
    "flowSystemMetadata": {
      "swaggerOperationId": "PostMessageToConversation"
    }
  }
}

```

Рис. 6. JSON схема триггера

ОСНОВНАЯ ЧАСТЬ

Приведем общую схему Flow получения динамических данных из хранилища SharePoint (рис. 1).

Семантику запроса можно обозначить в нотациях API (рис. 2).

Получение данных на уровне SharePoint можно задекларировать с наличием конструкции IF(), которая будет меняться при вызове потока. Постоянным значением в данном блоке остается адрес тенанта, на котором развернут стенд (рис. 3).

Значение, переданное в этот входной параметр потока, игнорируется, если входной параметр odata.

nextLink, переданный в поток, имеет значение, отличное от false. Значение этого входного параметра должно быть в следующем формате: {siteId}/lists/{listId}/root. Чтобы создать значение этого параметра для передачи в поток из PowerApps для любой заданной библиотеки документов на сайте, отображаемом в наборе результатов выполнения потока PA_Graph_Site_List, добавьте / root к значению SiteIdListLibraryId, как определено на шаге «Выбор возвращаемых полей» указанного потока. Единственное отличие между получением списка файлов и папок из корневой папки библиотеки документов и получением списка файлов и папок с любой заданной папкой, содержащейся в библиотеке документов, заключается в том, что для корневой папки библиотеки документов необходимо изменить путь / root к значению SiteIdListLibraryId, тогда как для папки

вы вместо этого шага добавляется значение/ items / {itemId}, где {itemId} представляет этот идентификатор папки [10,11]. Идентификатор папки должен быть представлен в предыдущем прогоне потока того же потока, который определен на этапе выбора возвращаемых полей потока, когда запрашивались файлы и суппапки родительской папки текущего элемента. На шаге выбора возвращаемых полей в потоке значением поля Id будет {itemId}, а значением поля ContentType для этой строки будет Folder. Выборка полей метаданных для возвращаемых значений (рис. 4).

Наиболее значимые свойства полей, которые использует общее решение, представленное в этой статье, что, в свою очередь, позволило авторам реализовать полностью динамический облегченный пользовательский интерфейс, имитирующий подмножество функций, которые в противном случае были бы возможны только с использованием встроенного SharePoint включают (таб. 1):

Результат исполненного процесса можно изучить в low-code клиенте Power platform (рис. 5):

Данный пример ответа получен при выполнении операции получения поэлементно единицы массива из источника списка, где основной тип контента, кото-

рый встречается на практике наиболее часто — это список (List ContentType). Список на платформе SharePoint или Dynamics 365 — это колоночное хранилище типизированных примитивов (текстовые строки, время, валюты, пользователи и другое), в которых по умолчанию нет возможности использования бинарных файлов, так как бинарные файлы как элементы решения располагаются в схожем типе хранилища данных, но имеют другой тип контента. Данная особенность связана с системой хранения информации на СУБД, где бинарный файл преобразуется в значение BLOB формата. В нашем примере мы использовали динамическое нахождение связи между двумя типами данных: реляционного представления метаинформации и бинарных файлов. Сложность этого шага связана с примером полезной нагрузки, которую авторы используют для создания схемы, необходимой для этого действия. Код такого взаимодействия через триггеры приведен ниже (рис. 6):

Заключение

В данной статье приведена реализации организации ETL процесса по получению документов и метаданных с динамически изменяемым источником данных. Данный метод может быть использован для других источников данных, который поддерживают интеграцию с платформой Microsoft 365.

ЛИТЕРАТУРА

1. Rezaee Jordehi A. Optimal placement of battery swap stations in microgrids with micro pumped hydro storage systems, photovoltaic, wind and geothermal distributed generators/ A. Rezaee Jordehi, Mohammad Sadegh Javadi, João P.S. Catalão //International Journal of Electrical Power & Energy Systems Volume 125, February 2021, 106483.
2. Gisbert R.S. Low-Code as Enabler of Digital Transformation in Manufacturing Industry/ R.S. Gisbert, O. Garcia, F. Fraile // Appl. Sci. 2020, 10(1), 12; <https://doi.org/10.3390/app10010012>.
3. How M.L. Artificial Intelligence for Social Good in Responsible Global Citizenship Education: An Inclusive Democratized Low-Code Approach / M.L. How, Y.J. Chan, S.M. Cheah, A.C. Khor // 19–21 of February 2021 in Vienna, Austria at the 2nd World Conference on Teaching and Education <https://doi.org/10.6084/m9.figshare.14038811.v1>.
4. Rauschmayr N. Amazon SageMaker Debugger: A system for real-time insights into machine learning model training [Electronic resource] / N. Rauschmayr, V. Kumar, R. Huligol and others //Amazon Science — URL: <https://www.amazon.science/publications/amazon-sagemaker-debugger-a-system-for-real-time-insights-into-machine-learning-model-training> (accessed:01.04.2021).
5. Straschnov E. You Shouldn't Have to Learn How to Code [Electronic resource] / E. Straschnov // Веб-страница—URL: https://www.huffpost.com/entry/you-shouldnt-have-to-lear_b_6111914 (accessed:01.04.2021).
6. Floridi L. GPT-3: Its Nature, Scope, Limits, and Consequences / L. Floridi, M. Chiriatti // Minds and Machines — 2020 — № 30. — P. 681–694.
7. Balasingham R. Designing a Data Warehouse System for Sales and Distribution Company / R. Balasingham, R. Subash // Big Data Mining and Analytics — 2021 — № 1 (6). — P. 1–7.
8. Dipanker J. Salesforce Integration Architecture/J.Dipanker, J. Hutcherson.— Apress, 2021.— 383 p.
9. Mccarthy S. Reusing Dynamic Data Marts for Query Management in an On-Demand ETL Architecture / S. Mccarthy // PhD thesis, Dublin City University. — 2021.
10. Voronkov I.A. Usage of collaborative filtering in sharing platforms / I.A. Voronkov, S.E. Saradgishvili //Journal of Physics: Conference Series — № 1679 (3). — P. 1–6.
11. Voronkov, I.A., Saradgishvili, S.E. Usage of a BART algorithm and cognitive services to research collaboration platforms // Journal of Physics: Conference Series, 2020, 1694(1), 012028.