

ERP or “enterprise resource planning” software helps automate many core business areas, such as procurement, production, materials management, sales, marketing, finance, and human resources. These data sources play a pivotal role in enterprise digital transformation initiatives. SAP is the market leader in ERP software. This whitepaper explains how DeepIQ’s DataStudio enables cross functional analytics by simplifying the integration of your SAP data with the rest of your data ecosystem.

SAP was one of the first companies to develop and market standard software for business solutions successfully and continues offering industry-leading ERP solutions. Solving complex problems in manufacturing, energy, mining, and other industries require ERP

data. These include:

- Root Cause Analysis: with early indications about potential hazards, you can take timely measures to avert problems
- Customized Insights: gain insights and improve the accuracy of processes, customers, and workflows
- Production Optimization: improve yields and production volumes by optimizing operational profiles of plants/equipment
- Production Quality: determine which processes, workflows, and factors contribute to quality objectives.

Combining Operational Data (SCADA and other time series sources) with ERP data will ensure machine learning models are complete and accurate. Traditionally, ingesting ERP data along with time series, geospatial, unstructured, and structured data has been manual and error prone. DeepIQ's DataStudio has built-in connectors to ingest SAP data into your cloud platform of choice and export data from your cloud platforms into SAP. This whitepaper explains how DeepIQ DataStudio provides multiple ways to integrate your SAP data with the rest of your cloud ecosystem.

DeepIQ is a self-service {Data + AI} Ops app built for the industrial world. DeepIQ simplifies industrial analytics by automating the following three tasks:

1. Ingesting time series, structured and geospatial data at scale into your cloud platform.
2. Implementing sophisticated time series and geospatial data engineering workflows; and
3. Building state-of-the-art ML models using these datasets.

In addition to reading directly from the underlying database, DeepIQ's DataStudio supports three modes of integration with SAP:

- IDocs
- BAPI
- OData

## IDocs

IDoc (Intermediate Document) is a standard data structure used in SAP applications to transfer data or information from SAP to other systems and vice versa. Outbound IDocs generated in SAP can be saved as XML files in the cloud or a different storage location.

DataStudio can read these XML files, parse them, and create flattened structures required for persistence in your cloud databases. As IDocs are processed asynchronously, they can be used to read and write an enormous number of transactions without overloading the SAP systems.

In the example below, we parse Material IDocs and write to the data lake of your choice.



The screenshot displays the Databricks Canvas environment. On the left, a sidebar contains navigation icons for Studio, Data, Jobs, and Admin. The main workspace shows a workflow with two components: 'Read SOAP' (Data Transformation) and 'Write Data Lake' (Data Transformation), connected by a green line. The 'Read SOAP' component is highlighted with a yellow border. The right sidebar shows the 'PROPERTIES' tab for the selected component, with fields for Source, Azure File Path, WSDL Endpoint, Headers, Username, Password, Bindings, and Input Type.

**Canvas**  
Select an experience to get started  
Spark

**COMPONENTS** **WORKFLOWS**

Try 'Workflow'

Workflows Starred  
Sort [icon] [icon] + New

Directories > SAP

- Soap Read
- OData - Read Equipment
- IDocs - read products

1 to 3 of 3 < >

**Read SOAP**  
Data Transformation

**Write Data Lake**  
Data Transformation

**PROPERTIES** **NOTES**

Description  
Read SOAP reads response from a SOAP web service.

Source  
Azure Blob Storage

Azure File Path  
wasba://devdatastudio@devdeepiglab.blob.core.winc

WSDL Endpoint  
http://erpukpitm.gotdns.org:8015/sap/bc/srt/rfc/sap/zbaf

Headers

Key	Value
soapAction	urn:sap-com:document
Content-Type	text/xml; charset=utf-8

Username  
EHPUSER498

Password  
deepig@123

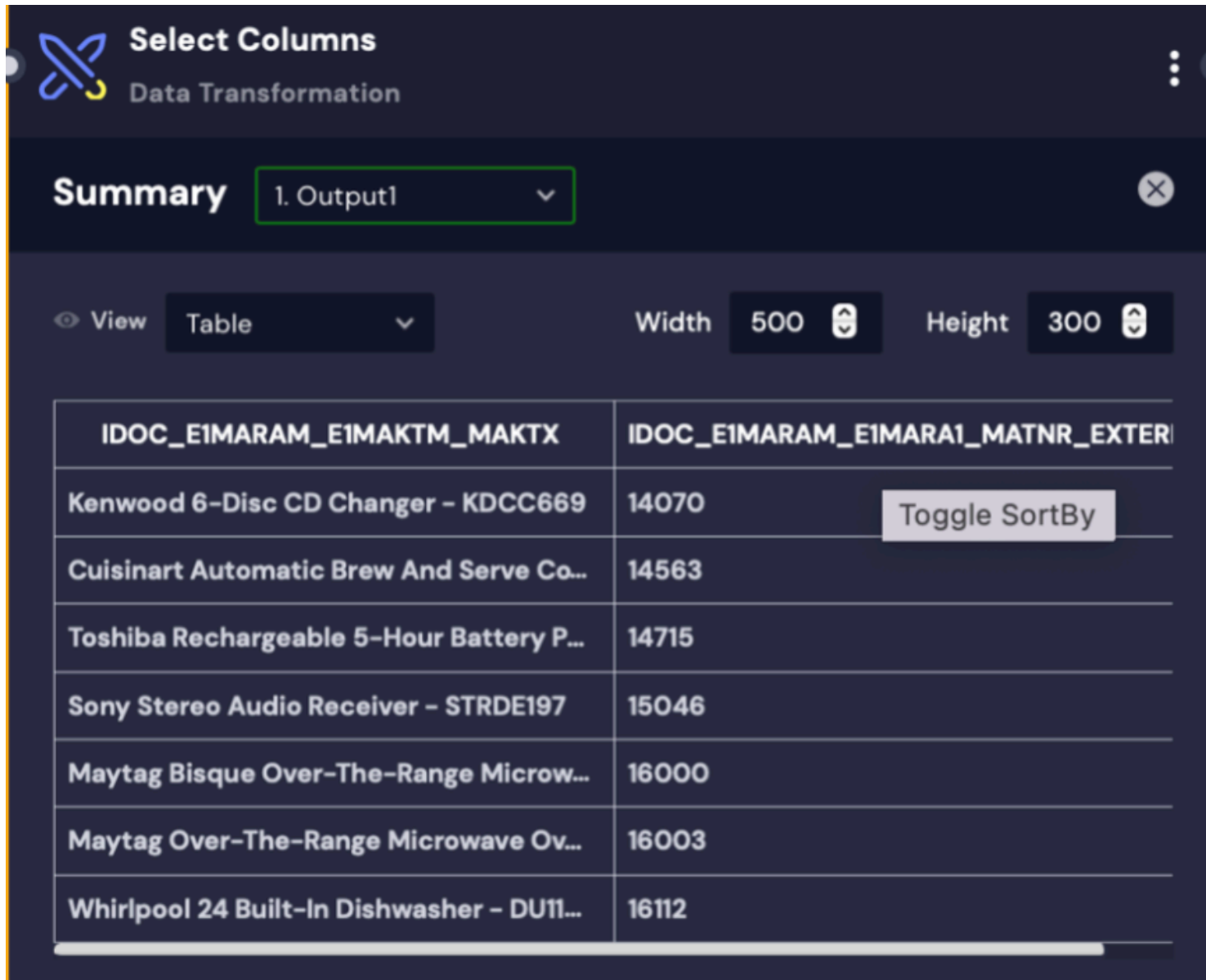
Bindings

Input Type

The following is an example of the Simple Object Access Protocol (SOAP) message that is received from SAP.

```
<?xml version="1.0"?>
<MATMAS05>
...
  <IDOC BEGIN="1">
    <EDI_DC40 SEGMENT="1">
      <TABNAM><![CDATA[EDI_DC40]]></TABNAM>
      <MANDT>800</MANDT>
      <DOCNUM>0000000002856760</DOCNUM>
      <DOCREL>740</DOCREL>
      <STATUS>03</STATUS>
      <DIRECT>1</DIRECT>
      <OUTMOD>2</OUTMOD>
      <IDOCTYP>MATMAS05</IDOCTYP>
      <MESTYP>MATMAS</MESTYP>
      <SNDPOR>SAPTST</SNDPOR>
      <SNDPRT>LS</SNDPRT>
      <SNDPRN>SEN1</SNDPRN>
      <RCVPOR>A000000068</RCVPOR>
      <RCVPRT>LS</RCVPRT>
      <RCVPRN><![CDATA[EDI_MM01]]></RCVPRN>
      <CREDAT>20220819</CREDAT>
      <CRETIM>044115</CRETIM>
      <SERIAL>20220819044115</SERIAL>
    </EDI_DC40>
    <E1MARAM SEGMENT="1">
      <MSGFN>009</MSGFN>
      <MATNR>ODATA9</MATNR>
      <ERSDA>20220819</ERSDA>
      <ERNAM>EHPUSER499</ERNAM>
      <LAEDA>00000000</LAEDA>
      <PSTAT>K</PSTAT>
      <MTART>FERT</MTART>
      <MBRSH>M</MBRSH>
      <MEINS>EA</MEINS>
      <BLANZ>000</BLANZ>
      <BRGEW>0.000</BRGEW>
      <NTGEW>0.000</NTGEW>
      <VOLUM>0.000</VOLUM>
```

DeepIQ's DataStudio converts the above SOAP message into an analytics ready dataset as shown below.



IDOC_EIMARAM_E1MAKTM_MAKTX	IDOC_EIMARAM_E1MARA1_MATNR_EXTERI
Kenwood 6-Disc CD Changer – KDCC669	14070
Cuisinart Automatic Brew And Serve Co...	14563
Toshiba Rechargeable 5-Hour Battery P...	14715
Sony Stereo Audio Receiver – STRDE197	15046
Maytag Bisque Over-The-Range Microw...	16000
Maytag Over-The-Range Microwave Ov...	16003
Whirlpool 24 Built-In Dishwasher – DU11...	16112

Using the IDoc approach has two advantages. Impact on the SAP server is minimal, and exporting data from SAP is relatively simple.

## BAPIs

SAP's BAPIs (Business Application Programming Interface) are methods for SAP business objects. These objects are stored in the Business Object Repository (BOR) of the SAP system and are used for carrying out business tasks.

BAPIs have standard business interfaces which enable DeepIQ's DataStudio to access SAP processes, functions, and data. DataStudio uses BAPIs to use the synchronous communication method to return data. SAP's BAPI uses the standard SOAP protocol to expose methods.

DataStudio can read SOAP Webservices that are exposed by SAP.

- The 'Read BAPI' component parses the WSDL file generated by SAP and allows the user to choose the binding to be used.
- Once the binding is selected, a sample request XML is generated, which the user can modify.
- Users can add the authentication credentials and run the workflow.

Thus, the 'Read SOAP' component makes the consumption of complex SAP web services straightforward. The following image shows DeepIQ's DataStudio ingesting SAP data through the BAPI interface.



The screenshot shows the DataStudio Canvas interface. On the left, the 'COMPONENTS' sidebar is visible, with 'WORKFLOWS' selected. A search bar contains 'Try "Workflow"'. Below it, a list of workflows is shown, including 'IDocs - read orders', 'Soap Read', 'OData - Read Equipment', and 'IDocs - read products'. The 'Soap Read' workflow is highlighted. The main canvas area displays a workflow with two components: 'Read SOAP' and 'Write Data Lake'. The 'Read SOAP' component is highlighted with a yellow border and shows a table of equipment data. The 'Write Data Lake' component is connected to the 'Read SOAP' component via a data link. The 'Read SOAP' component's 'Summary' dropdown is set to '1. Output!'. The 'Write Data Lake' component's 'Summary' dropdown is set to '2. Summary!'. The 'Read SOAP' component's table has the following data:

DESCRIPT	EQUICATGRY	EQUIPMENT	HIGH	LOW
test	V	10001		
Test Number range 04	C	10002		
MRC 1	V	10003		
222	V	10004		
Sandero	V	10005		
TRATOR CATERPILAR CA 777F	V	10006		
TRATOR CATERPILAR CA 778F	V	10007		
Trator Caterpillar CA 778F	V	10008		
Caminhão Basculante Volvo 555	V	10009		

In this example, we read Equipment data from SAP and persist to the data lake.

The BAPI method is high performing, guarantees constancy across SAP versions, and is business object-oriented. BAPI can be configured to retrieve data on a schedule to ensure that the data in DataStudio is current. This method may have performance implications on the SAP server.

SAP OData is a standard Web protocol used for querying and updating data present in SAP. In SAP, we use a transaction code to create an OData Service. OData is used to build and consume RESTful APIs. DataStudio can read SAP OData Web services using the 'Read Rest' component.

As the SAP OData web services have a custom authentication mechanism, we use DataStudio's data transformation capabilities to create the HTTP request payloads, parse the JSON response and generate the data frame.

The following shows a workflow to ingest and process SAP Material data, including authentication. We can also create or update SAP data in a similar way.

DeepIQ is on a mission to transform industrial processes by digitizing industrial expertise. Our vision is to drive end-to-end automation, enabling systems such as self-running power plants or drilling rigs using generative AI as the higher order reasoning layer operating over existing industrial automation technology stack.

[Privacy Policy](#)   [Terms of Services](#)



FEATURES

Extract  
Engineer  
Explore

INDUSTRIES

Upstream  
Midstream  
Downstream & Chemical  
Mining

SOLUTIONS

IT-OT Contextualization  
Well Construction  
Optimization  
P & ID Digitization  
Predictive Maintenance  
Production Optimization  
Route Optimization  
Sustainability

PARTNERS

AWS  
Azure  
Cloudera  
Google  
Databricks  
OSIsoft(AVEVA)  
Snowflake

COMPANY

[About Us](#)  
[Resources](#)





News

DataStudio Deployment

Guide

Career

Contact Us

