SMTDATA
IT Enterprise Insights

Using the ITBI Data Lake

itbi
SMTDATA

SMTDATA
IT Enterprise Insights

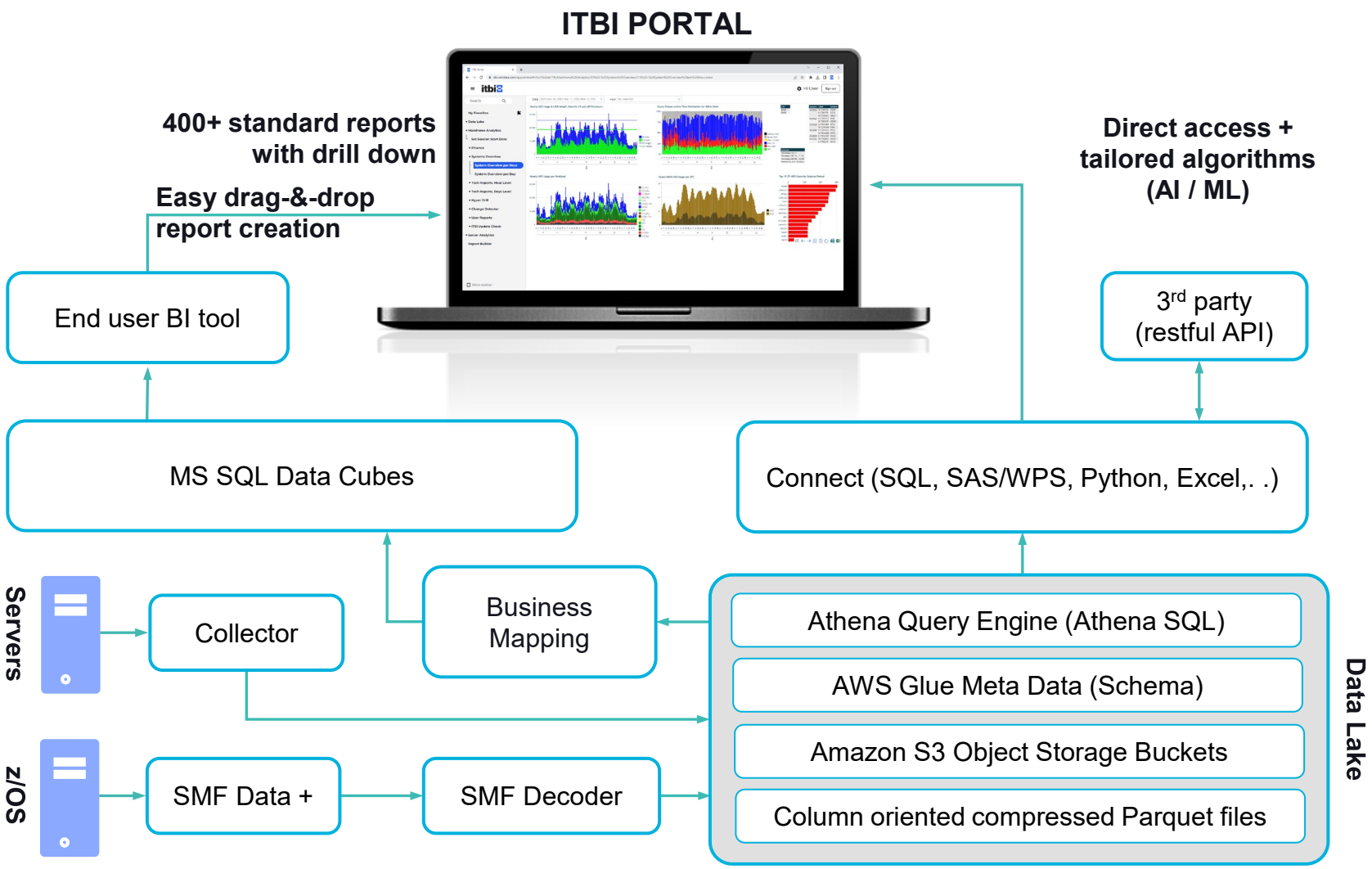
ITBI Architecture

What data is in the Data Lake?

How do you access it?

Examples

02 ITBI Architecture



SMTDATA
IT Enterprise Insights

Data in the Data Lake

Raw SMF in Tables:

- Essentially all fields from all SMF types we support
- Mostly decoded

MXG-like Tables:

- Only selected fields from the SMF types we support

These tables can be accessed via SQL from almost any tool (SAS, Python, Excel, Athena, R, and many more)

The tables are updated within minutes of the data arriving

SMTDATA
IT Enterprise Insights

Raw SMF Tables

Reporting on fields that are not supported in the cubes

Reporting on details that are aggregated away in the cubes

Reporting on ‘event’ based data

Reporting on recent data – within minutes of data being received by us

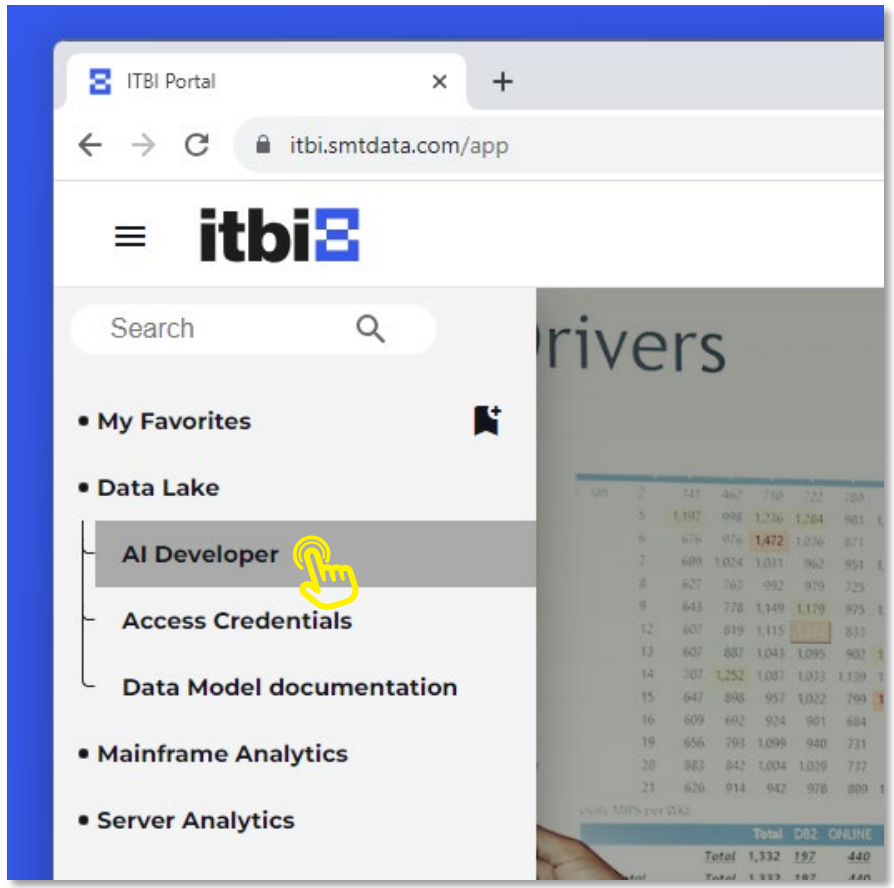
Requires a good understanding of SMF

MXG Like Tables

Reuse of existing SAS programs that work on MXG tables, but note, only selected fields are supported

Requires a good understanding of MXG

SMTDATA
IT Enterprise Insights

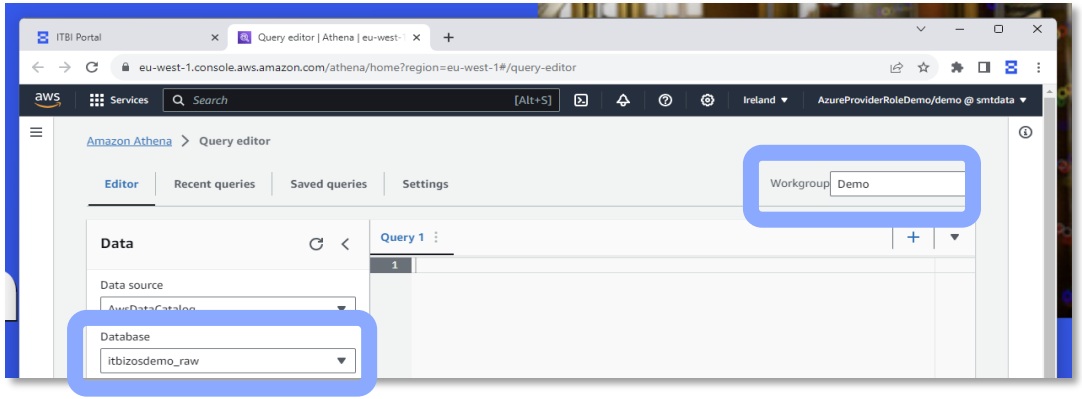


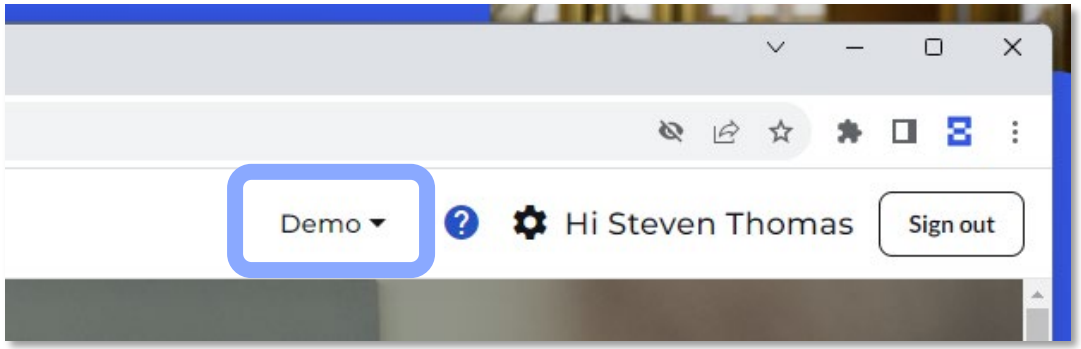
Access via SQL from Athena

Log on to the Portal

Choose Data Lake

Choose AI Developer





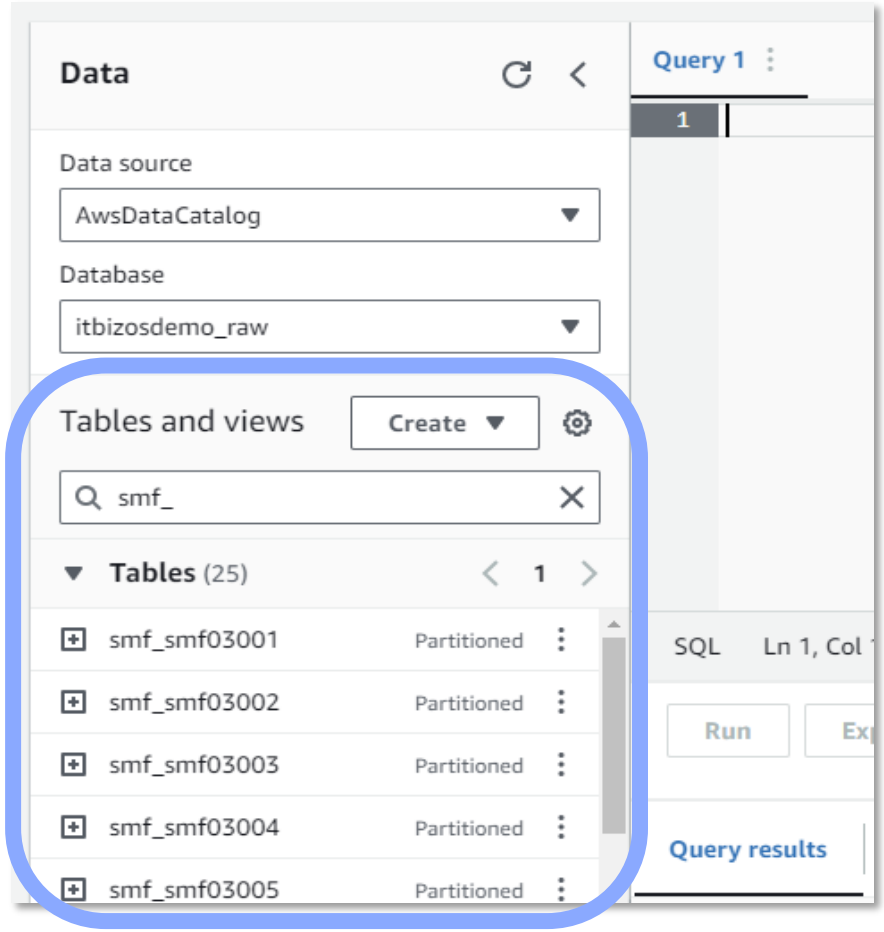
Access via SQL from Athena

Set the Workgroup to your company name

Set the Database to the ‘itbizos…_raw’ where ‘…’ is your company name

The company name should be spelled the same way it is in the Portal

SMTDATA
IT Enterprise Insights



Access via SQL from Athena

You can see a list of the relevant tables and views on the left side of the screen.

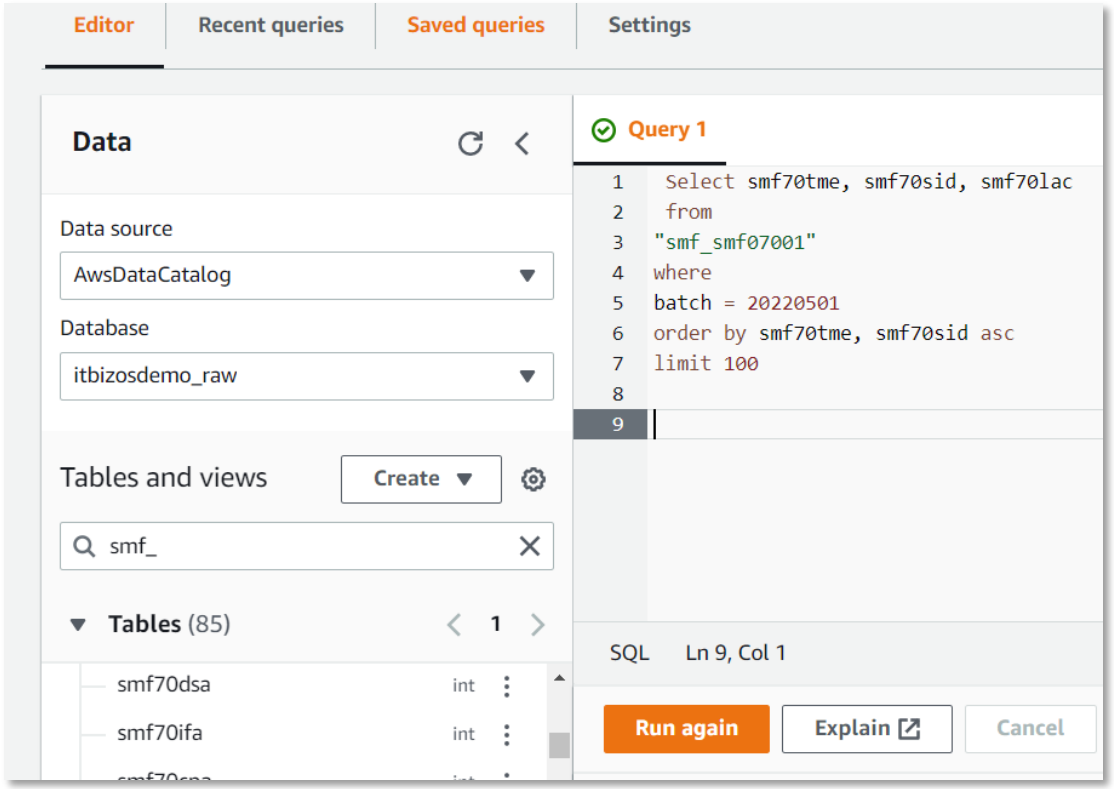
The raw SMF tables start with smf_*

- The main tables are named smf_smfyyyzz
- Where yyy is the SMF number
- And zz is the subtype
- So smf_smf03001 is SMF30 subtype 1

The mxg-like views start with v_smtmxg_*

- In general, the views follow the mxg naming conventions

SMTDATA
IT Enterprise Insights



This example shows

A select statement that chooses the following columns from the SMF70 subtype 1 table:

- SMF70tme (100^{ths} of a second since midnight)
- SMF70sid (LPAR system name)
- SMF70lac (4HRA MSU)

For the date 2022-05-01 (batch number is the date)

Sorted by time and system

Limited to 100 rows

SMTDATA
IT Enterprise Insights

Tips, tricks and documentation

You can cut and paste the results or download them as csv using Athena.

Always include a where clause specifying the date (or range of dates) of interest using ‘batch’. E.g. where batch > 20221001 will give you all days after October 1, 2022. If you don’t specify a batch, then Athena will scan all of the data in the data lake.

When developing a query or just exploring the data it is a good idea to limit the number of rows returned using the ‘limit’ statement. Then you can remove the limit when you are ready to ask for a full set of data.

SMTDATA
IT Enterprise Insights

Tips, tricks and documentation

We are developing additional documentation. Have a look here to see the version under development: https://dev-api.smtdata.com/data-model-documentation/swagger/index.html

You can also find lots of documentation of SMF on the IBM website: https://www.ibm.com/docs/en/zos/2.4.0?topic=smf-records

SMTDATA
IT Enterprise Insights

Thank you

Steven Thomas, CTO

steven.thomas@smtdata.com
+45 2056 8259
linkedin.com/in/stevendthomas
smtdata.com

Stay in touch by following our LinkedIn profile