What is the piRate, Snake & Cup of Coffee Doing in My Database

SQL Server Extensibility Framework & External Languages

Niels Berglund
niels.it.berglund@gmail.com
https://nielsberglund.com
@nielsberglund

PASS SQLSATURDAY CAPE TOWN | 14 SEP 2019

Thank Sponsors

Niels Obligatory Shameless Self Promo

- Software Architect - Derivco.
- Author - "First Look at SQL Server 2005 for Developers".
- Microsoft Data Platform MVP.
- Researcher / Instructor - DevelopMentor.
- Speaker - TechEd, DevWeek, SQL Pass, etc.
- Longtime user of SQL Server.
- Working closely with MS around SQL Server.

https://nielsberglund.com

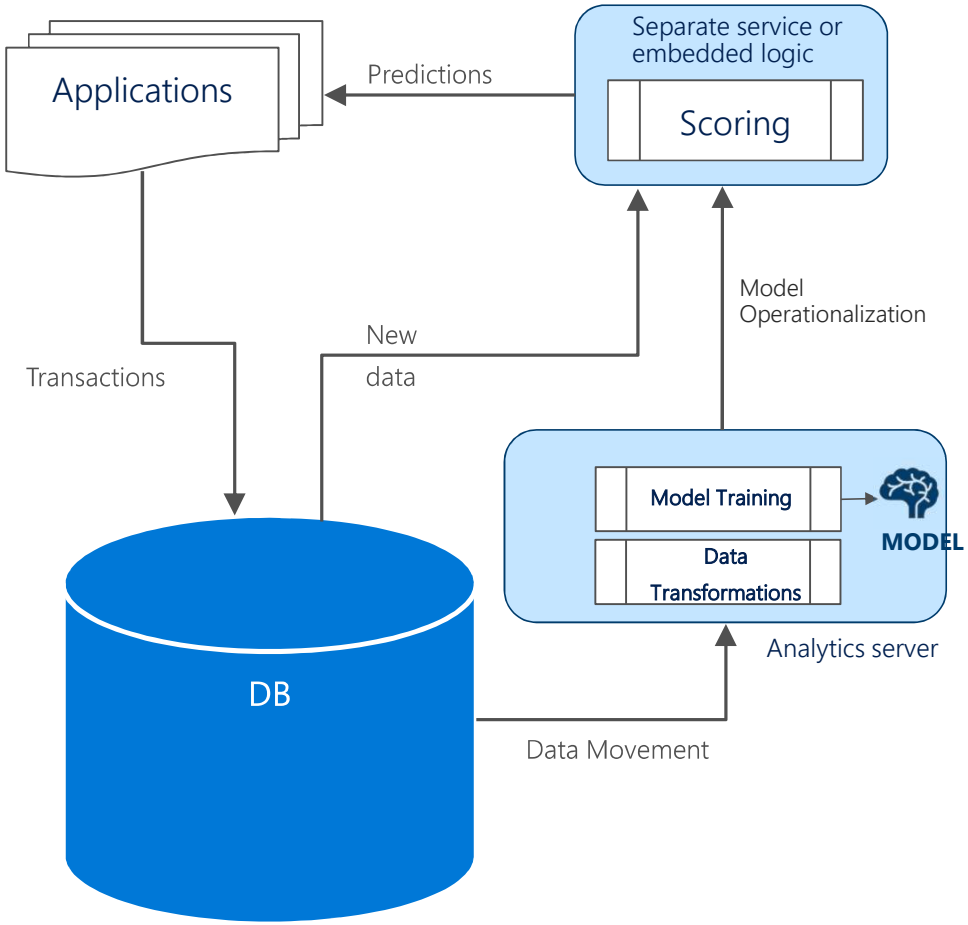
SQL Server Language History

- Dinosaur era (i.e. pre. SQL Server 2005):
  - SQL.
  - T-SQL.
  - Extended stored procedures (XP).
- SQL Server 2005:
  - SQLCLR!
- SQL Server 2016:
  - R
- SQL Server 2017:
  - Python

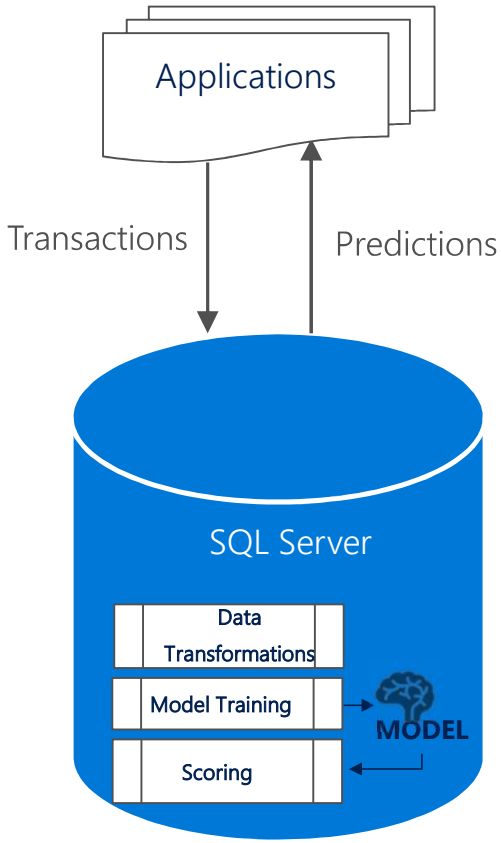
Why In-database ML with SQL Server?

- Better Collaboration and Insights Sharing
- Streamlined Deployment of R/Python Scripts and Models
- Faster Time to Insights
- Better Security and Compliance

Streamline Productivity and Simplify Deployment



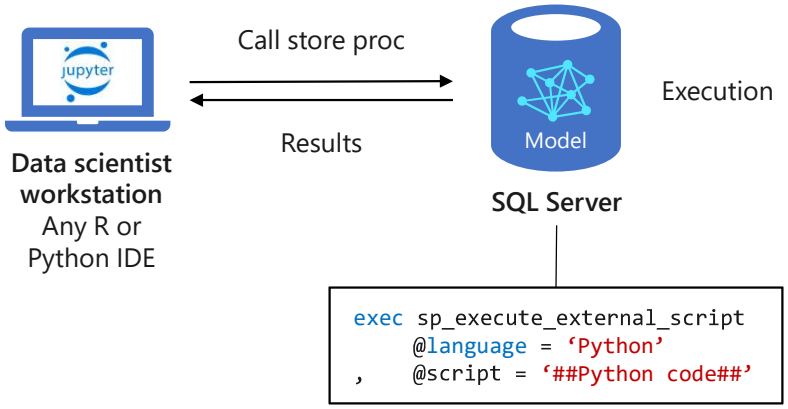
Machine Learning outside of DB



In-DB Machine Learning

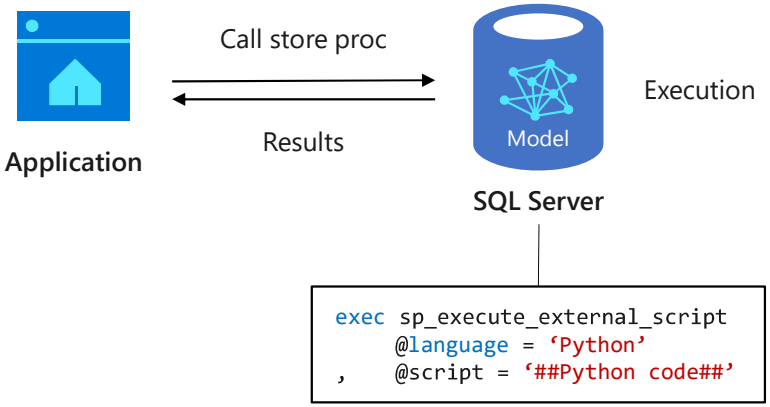
Data Scientists: Explore, Develop Models, Deploy

- Work with full datasets, no data movements.
- Use your favorite IDE, leverage power of SQL Server.
- Use any open source package in-db.
- Leverage scalable, fast MS algorithms.
- Streamlined model deployment.
- Model version management in-db.



Data and Application Developer - Model Consumption

- App developers: make applications intelligent:
  - Consume model by calling T-SQL procs.
  - No knowledge of models.
  - No conversions into other languages needed.
- Data developers: leverage the power of R/Python:
  - General purpose data processing.
  - Powerful data visualizations.



Run R & Python In-Db

```
1  EXEC sp_execute_external_script
2  @language =N'R',
3  @script=N'
4  OutputDataSet <- InputDataSet;
5  ',
6  @input_data_1 =N'SELECT 1 AS hello'
7  WITH RESULT SETS (([hello] int not null));
8  GO
9
```

RESULTS

	hello
1	1

```
1  EXEC sp_execute_external_script
2  @language =N'Python',
3  @script=N'
4  OutputDataSet = InputDataSet;
5  ',
6  @input_data_1 =N'SELECT 1 AS hello'
7  WITH RESULT SETS (([hello] int not null));
8  GO
9
```

RESULTS

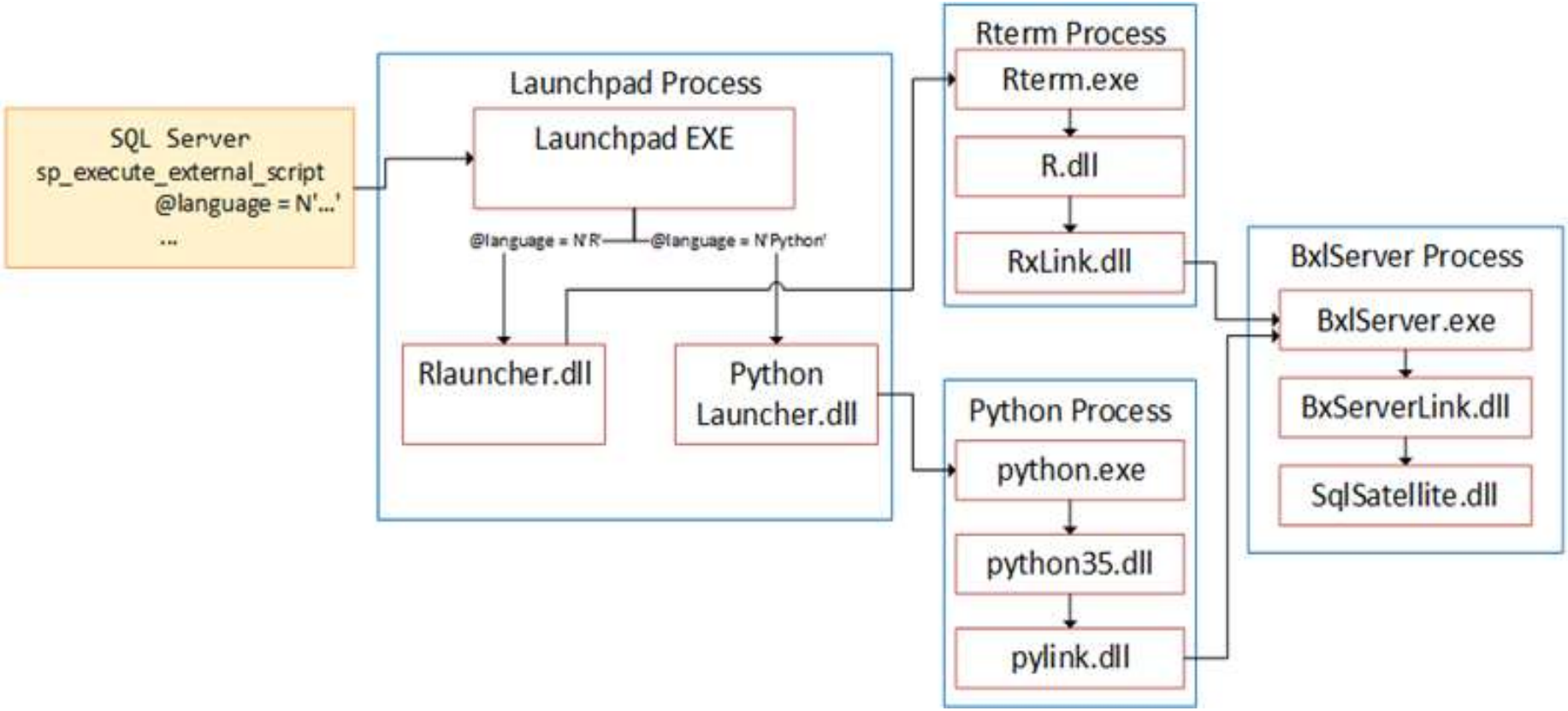
	hello
1	1

https://nielsberglund.com

External Languages

- R & Python called external languages.
- External to SQL Server.
- Does not run in the same memory space/process.
- Quite opposite to SQLCLR, similar to XP.

Architecture - I



https://nielsberglund.com

SQL Server Language History - II

- SQL Server 2017 (recap):
  - R & Python - SQL Server Machine Learning Services.
- SQL Server 2019:
  - Java
  - Wait a minute, did you say Java - Microsoft & Java?
- Why Java?

Why Java?

- Java big in Big Data (Spark etc.).
- More use cases than pure ML.
- Java compiled code - performance.
- Certain ML frameworks require Java (JVM).

Extensibility Framework

- Microsoft introduced the SQL Server Extensibility Framework.
- A way to incorporate arbitrary languages in SQL Server.
- C++ dll acts as a "bridge" between SQL Server and the application.
- The dll exposes well known interfaces which SQL Server calls into.
- The dll loads the external language runtime and hosts the application.

Registering an External Language

- To use an external language, its dll needs to be registered.
  - Any language except for R/Python.
- The language is registered per database in an instance.
- You register the language with CREATE EXTERNAL LANGUAGE.

```
CREATE EXTERNAL LANGUAGE <lang_name>
[ AUTHORIZATION <owner_name> ]
FROM (CONTENT = { <file_path> | <content_bits>,
      FILE_NAME = <external_lang_file_name>
    [ , PLATFORM = <platform> ]
    [ , PARAMETERS = <external_lang_parameters> ]
    [ , ENVIRONMENT_VARIABLES = <external_lang_env_variables
> ] )
[;]

-- SELECT * FROM sys.external_languages
```

Using an External Language

- You have registered an external language - now what?
- You place your code in a well known place, or in the database.
  - If well known place, the language dll need to understand where to load the code from.
- To place the code in the database you use `CREATE EXTERNAL LIBRARY`.

```
CREATE EXTERNAL LIBRARY <library_name>
[ AUTHORIZATION <owner_name> ]
FROM (CONTENT = { <file_path> | <library_bits> }
      [, PLATFORM = <platform> ])
WITH ( LANGUAGE = <language> )
[ ; ]

-- SELECT * FROM sys.external_libraries
```

Calling External Language Methods / Scripts

- Entry point is `sp_execute_external_script`.
  - Regardless if it is scripting language, or compiled language.
- The language extension must know what to do.

```
EXECUTE sp_execute_external_script
             @language = N'<language>',
             @script = N'<script> | <method>',
             [ @input_data_1 =  N'<input_data>' ]
             [ , @input_data_1_name = N'<input_data_1_name>' ]
             [ , @output_data_1_name = N'<output_data_1_name>' ]
             [ , @parallel = 0 | 1 ]
             [ , @params = ] N'@parameter_name data_type [ OUT | OUTPUT ] [ ,...n ]'
             [ , @parameter1 = ] 'value1' [ OUT | OUTPUT ] [ ,...n ]
[ WITH <execute_option> ]
[;]
```

Examples

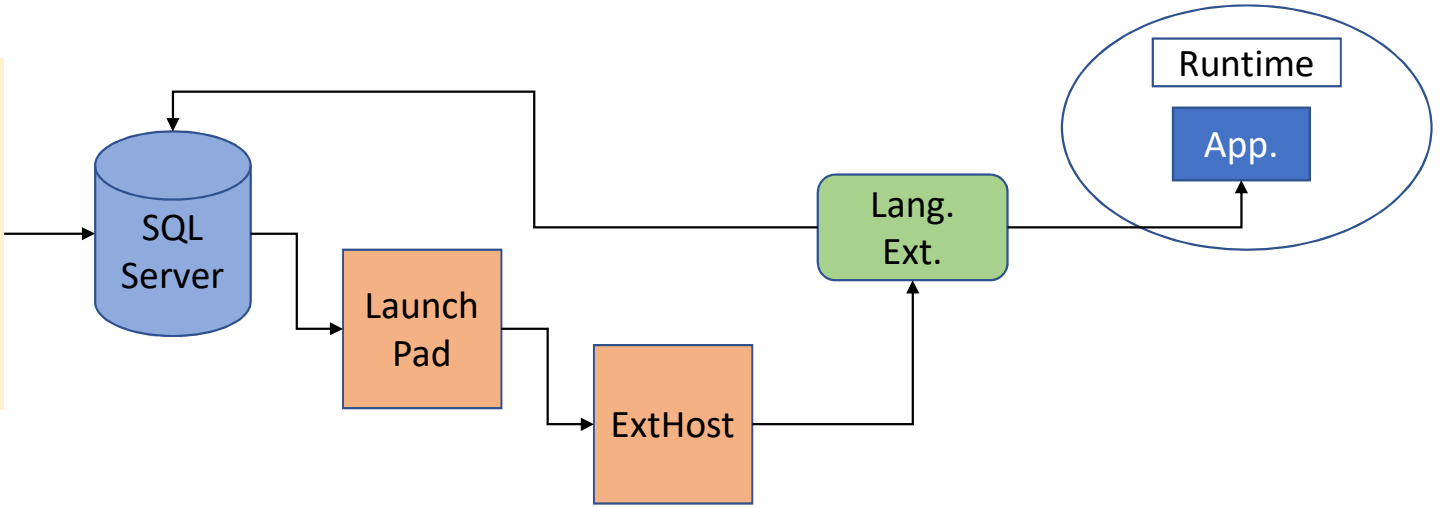
```
EXEC sp_execute_external_script  @language =N'R',
                                 @script=N'OutputDataSet<-InputDataSet',
                                 @input_data_1 =N'SELECT 42'
WITH RESULT SETS (([TheAnswer from R] int not null));
```

```
EXEC sp_execute_external_script  @language =N'Python',
                                 @script=N'OutputDataSet = InputDataSet',
                                 @input_data_1 =N'SELECT 42'
WITH RESULT SETS (([TheAnswer from Python] int not null));
```

```
DECLARE @p1 int = 21;
DECLARE @p2 int = 21;
EXEC sp_execute_external_script
  @language = N'Java'
, @script = N'sql.JavaTest1'
, @params = N'@x int, @y int'
, @x = @p1
, @y = @p2
GO
```

Architecture - II

```
DECLARE @p1 int = 21;
DECLARE @p2 int = 21;
EXEC sp_execute_external_script
  @language = N'Java'
, @script = N'sql.JavaTest1'
, @params = N'@x int, @y int'
, @x = @p1
, @y = @p2
GO
```



Java Code

- Code needs to be written in a specific way.
  - Class needs to inherit a class from an MS developed SDK.
- The class has an execute method which is called from the language extension.
- Helper classes for parameters and datasets.

Java Code Example

```
package sql;
import com.microsoft.sqlserver.javalangextension.*;
import java.util.LinkedHashMap;

public class JavaTest1 extends AbstractSqlServerExtensionExecutor {

public JavaTest1() {
  executorExtensionVersion = SQLSERVER_JAVA_LANG_EXTENSION_V1;
  executorInputDatasetClassName = PrimitiveDataset.class.getName();
  executorOutputDatasetClassName = PrimitiveDataset.class.getName();
}

public PrimitiveDataset execute(PrimitiveDataset input,
                        LinkedHashMap<String, Object> params) {

int x = (int)params.get("x");
int y = (int)params.get("y");

System.out.printf("The result of adding %d and %d = %d", x, y, x + y);
return null;
  }
}
```

```
DECLARE @p1 int = 21;
DECLARE @p2 int = 21;
EXEC sp_execute_external_script
  @language = N'Java'
, @script = N'sql.JavaTest1'
, @params = N'@x int, @y int'
, @x = @p1
, @y = @p2
GO
```

Deploy Java Code

- Deploy via `CREATE EXTERNAL LIBRARY`.
- Best practice to deploy .jar file.
- Need to deploy SDK as well.

```
-- deploy Java SDK
CREATE EXTERNAL LIBRARY javaSDK
FROM (CONTENT = 'W:\mssql-java-lang-extension.jar')
WITH (LANGUAGE = 'Java');
GO
```

```
-- deploy Java App
CREATE EXTERNAL LIBRARY mySqlJar2
FROM (CONTENT = 'W:\sql-1.0.jar')
WITH (LANGUAGE = 'Java');
GO
```

Future

- In SQL Server 2019:
  - R
  - Python
  - Java
  - More languages are on the way.
- Microsoft will open source the extensibility framework.
  - The Java SDK already open source.

Summary

- External languages:
  - R
  - Python
  - Java
- Code executes outside of SQL Server memory, and process space.
- Extensibility framework allows you to enable other languages.
- To enable a language a C++ extension dll is required.

PASS SQLSATURDAY CAPE TOWN | 14 SEP 2019

Thank Sponsors

Thank You!
Questions?

Niels Berglund
niels.it.berglund@gmail.com
https://nielsberglund.com
@nielsberglund