Data Science
DataOps
In the Microsoft Cloud

strategy insight engineering cloud

endjin

work smarter

ACCELERATE BUSINESS INNOVATION WITH THE MICROSOFT CLOUD

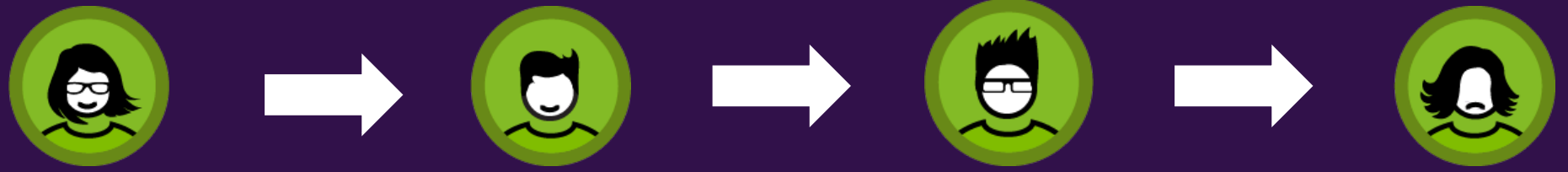
For EXAMPLE:

- Innovation and New Product Development Consultancy
- Application Proof of Concept & Prototyping
- Cloud Systems Architecture Guidance & Support

➡Big Data, Advanced Analytics & Visualization

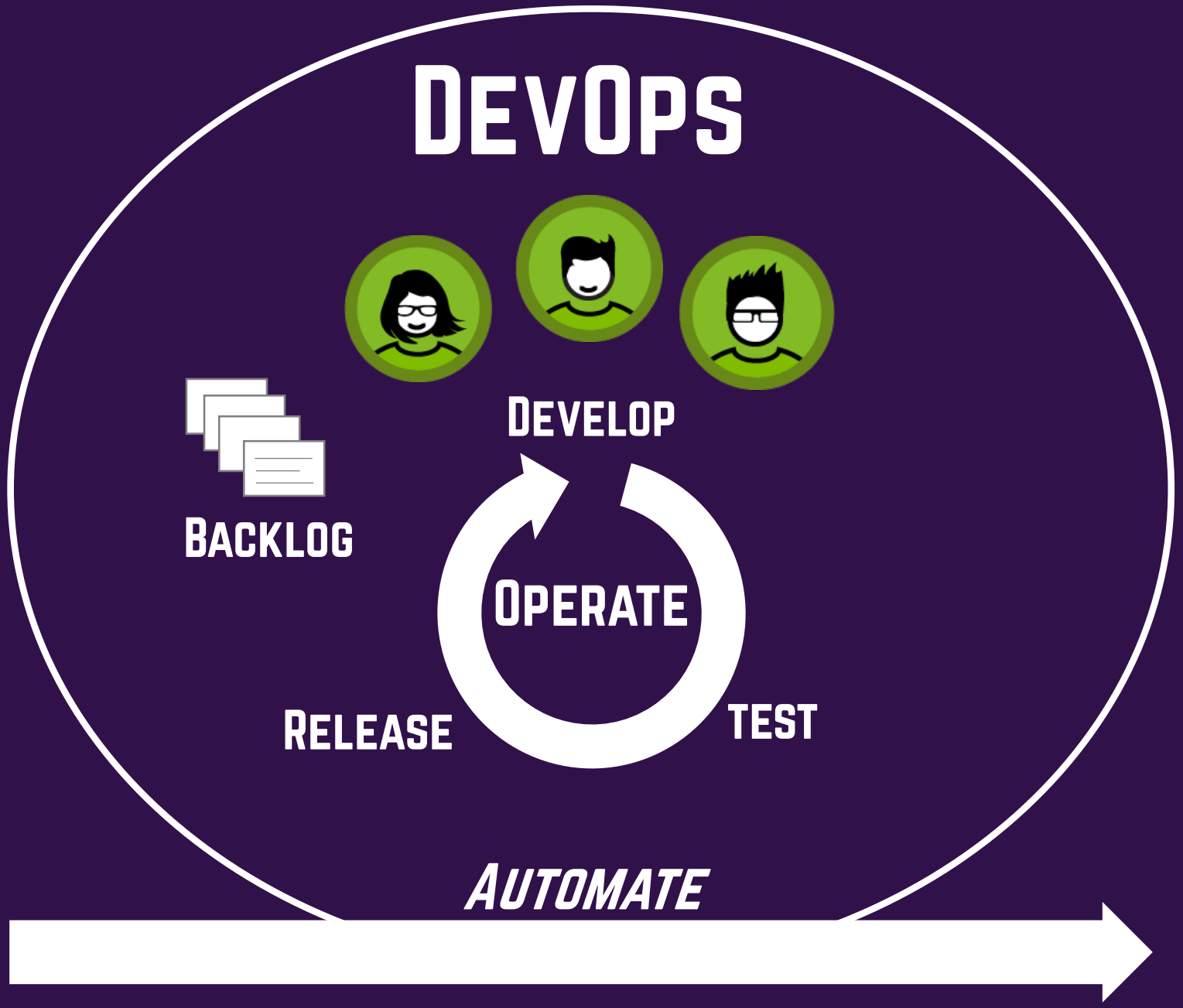
endjin

How it was...



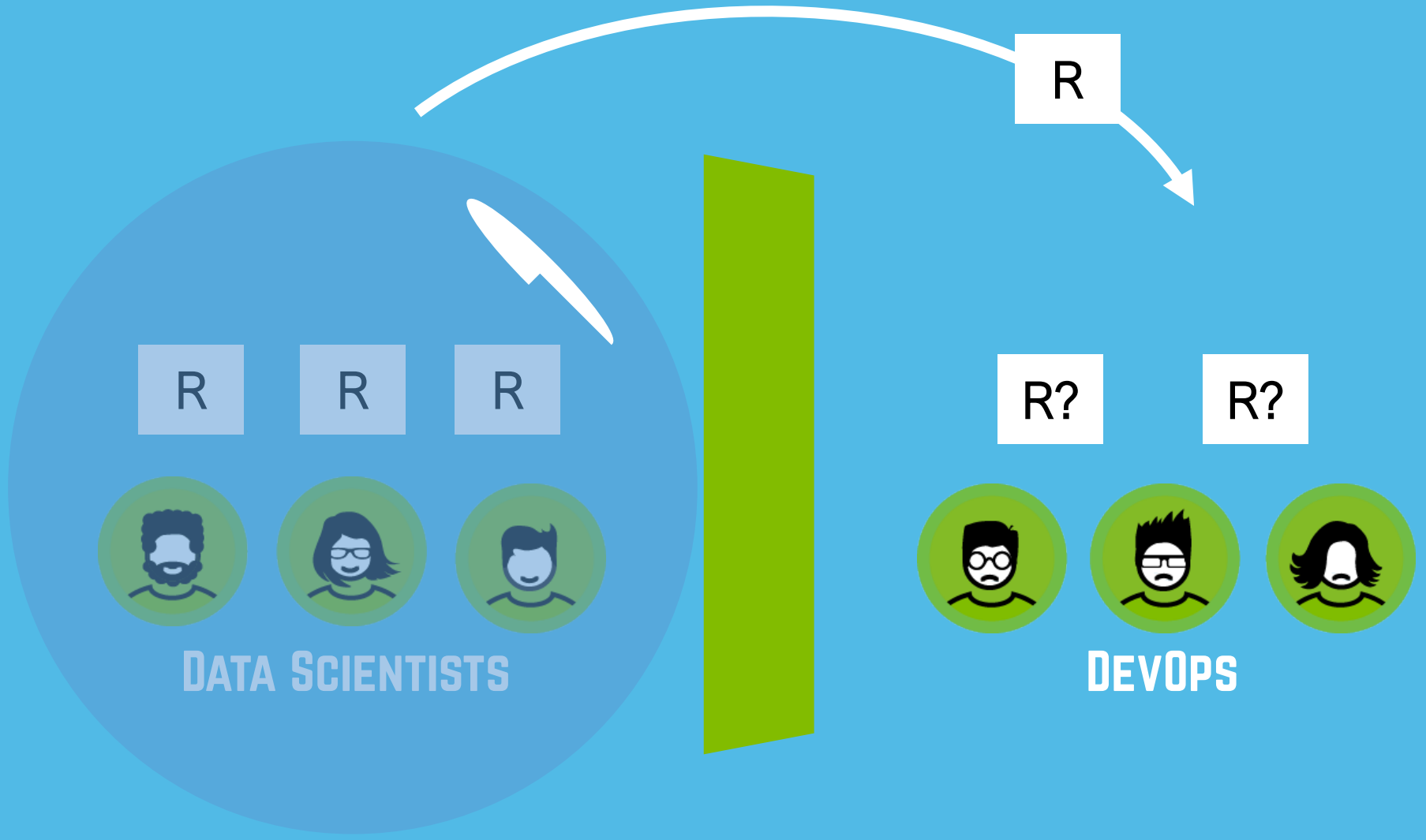
No feedback
Big releases
No accountability
Lack of Agility
Waste

Solved!



VSTS

The Data Science Problem today....



Backlog?
Version Control?
Quality?
Performance?
Release?
Auth?

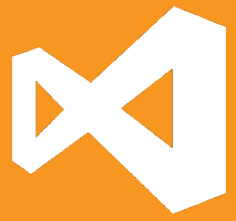
What if....

R
Development

Version Control
Test
Release

Operate

Azure Machine Learning
Microsoft R Server
R Tools for Visual Studio
R Studio



VSTS

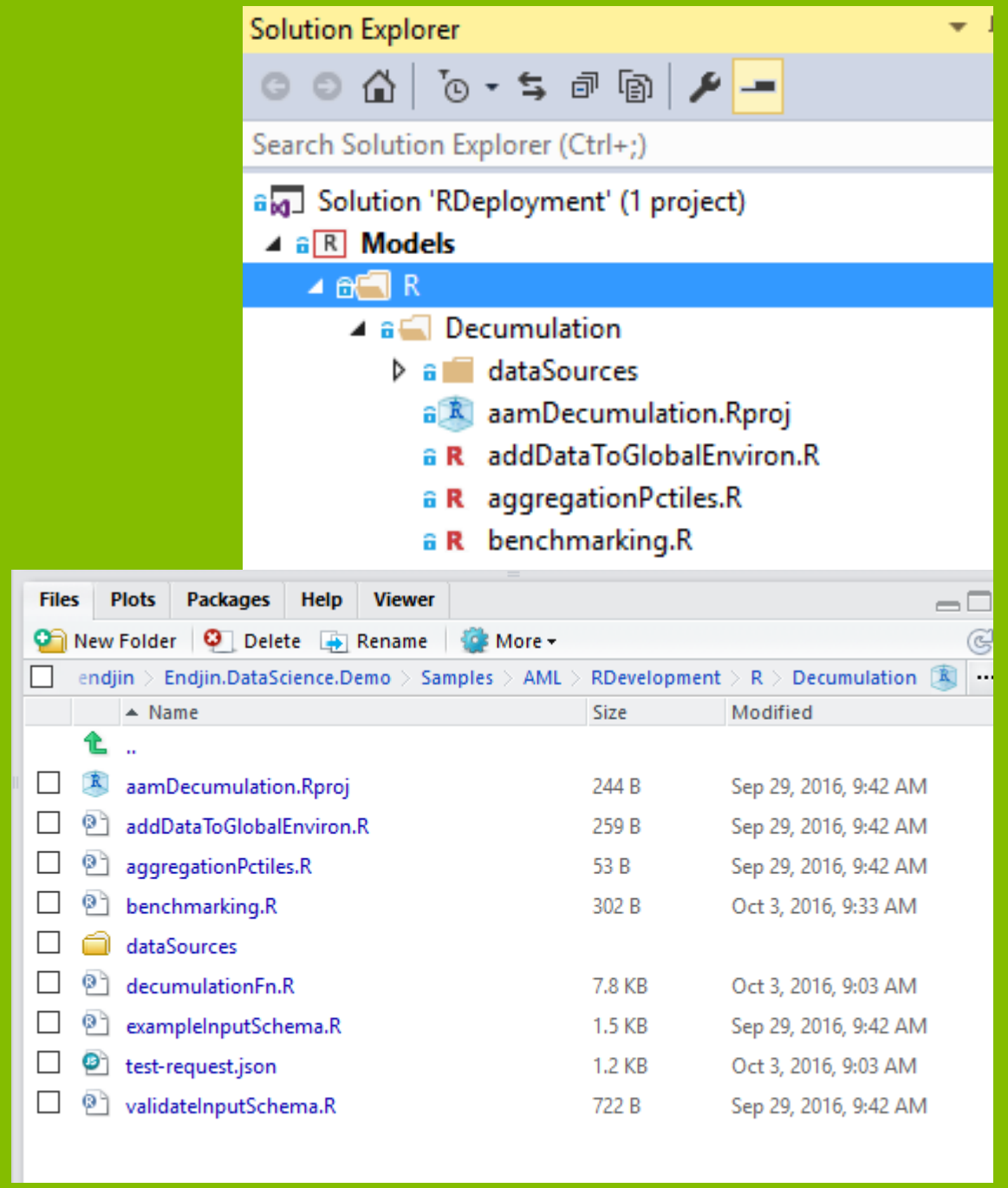
Azure Machine Learning
Deploy R
API Management
Active Directory

DEMO

Automated R deployments with VSTS:
https://vimeo.com/159813974

Automated R unit tests with VSTS:
https://vimeo.com/167279100

Take an R model in R Studio or Visual Studio...



Add some tests...

test-sum.R

```
context("sum")

test_that("sum adds positive numbers correctly", {
    expect_equal(sum(1, 1), 2)
    expect_equal(sum(1, 2), 3)
    expect_equal(sum(2, 2), 4)
    expect_equal(sum(2, 3), 5)
})
```

Add a deployment definition file...

deploy.json

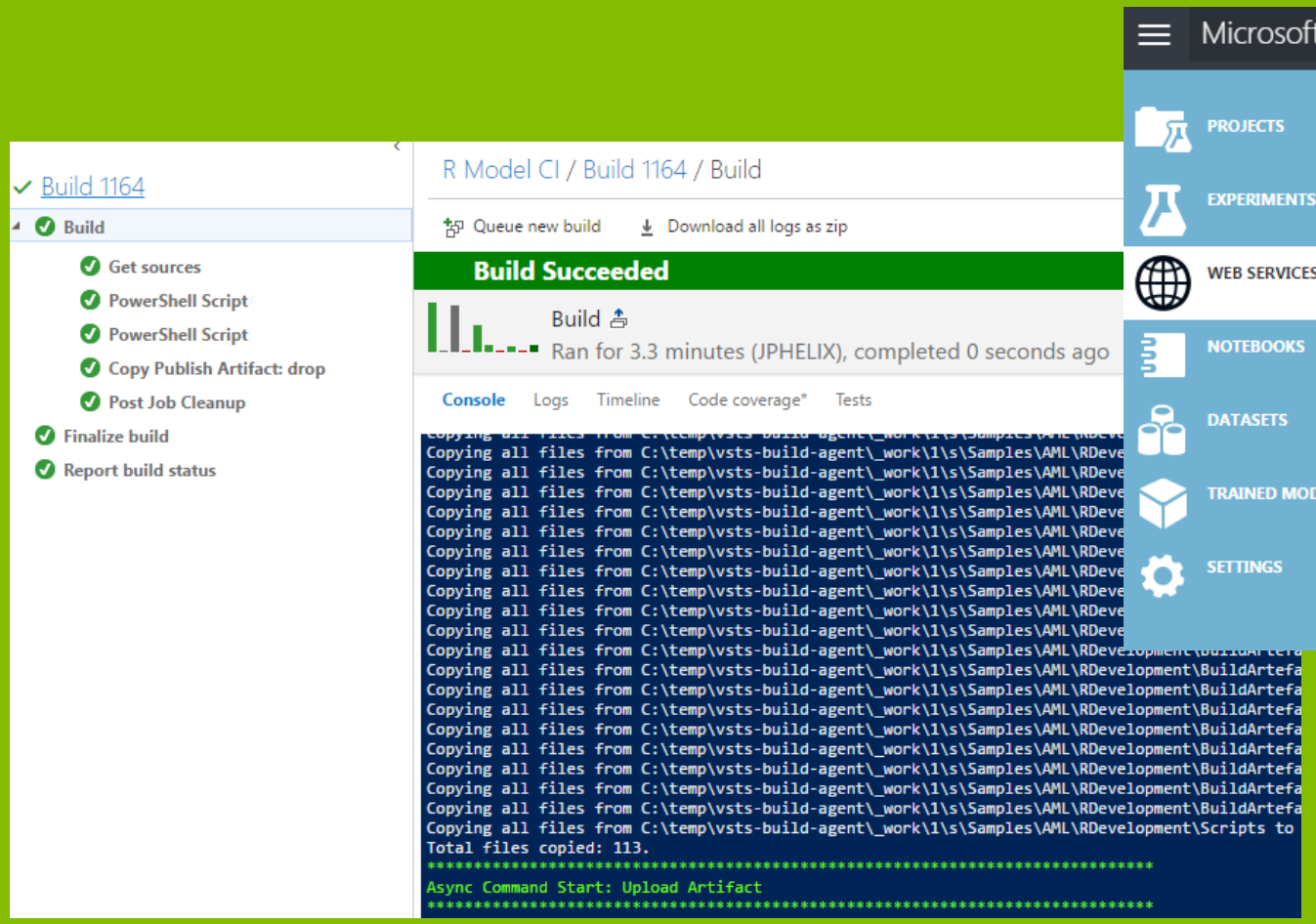
Schema: <No Schema Selected>

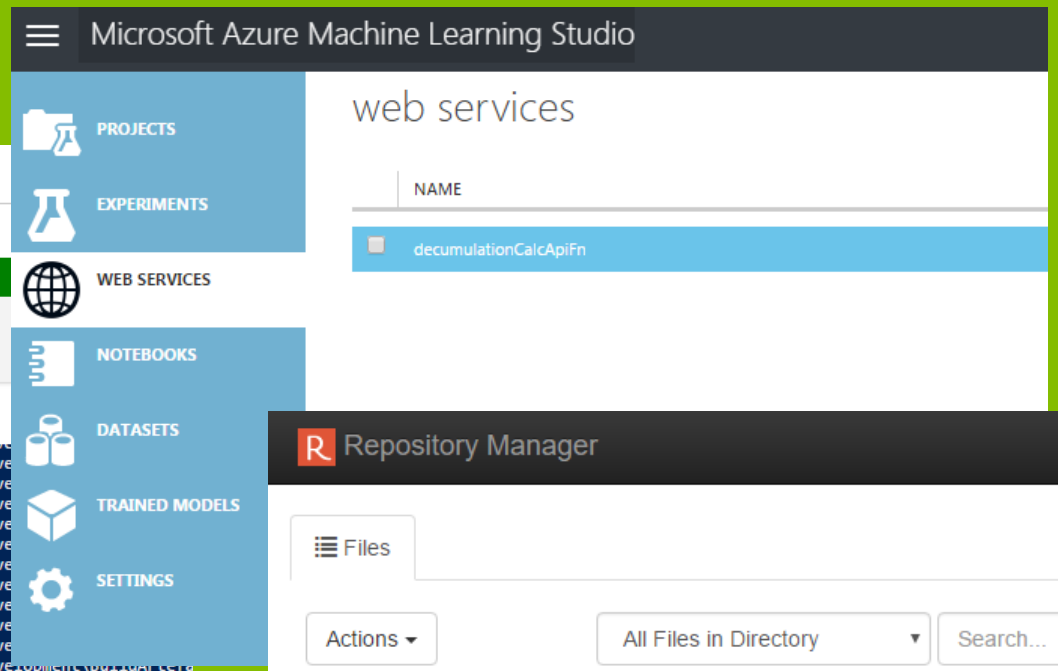
```
1  [
2    {
3      "script": "R/Decumulation/decumulationFn.R",
4      "function": "decumulationCalcApiFn",
5      "inputSchema": {
6        "input": "string"
7      },
8      "outputSchema": {
9        "result": "string"
10     },
11     "packages": [ "dplyr" ]
12   }
13 ]
```

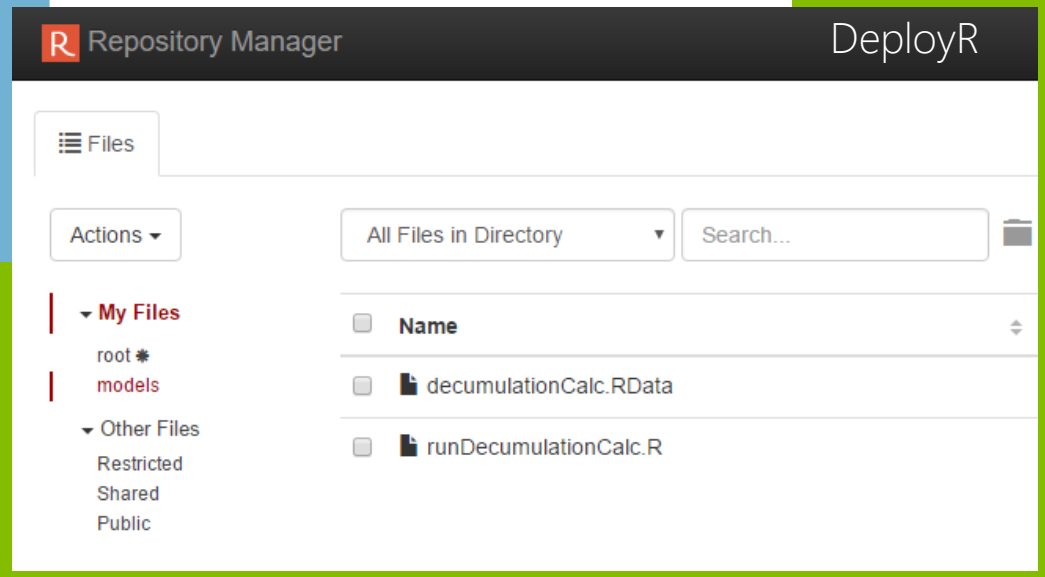
Add Model to version control...

```
> git add -A
> git commit -m "Adding model"
> git push
```

Continuous integration and deployment for R with VSTS...







Data Science DataOps

Project Management	Development	Quality Assurance	Release Management	Operations
Manage Priorities Work Item Tracking Progress Reporting	R Models R Packages AML Experiments	Unit Testing Model Validation Performance Testing	Model deployment Package Management Policies Audit	Hosting Scale Authentication Metering Throttling Monitoring

VSTS

+

Azure Machine Learning

DeployR

+

API Management

+

endjin