


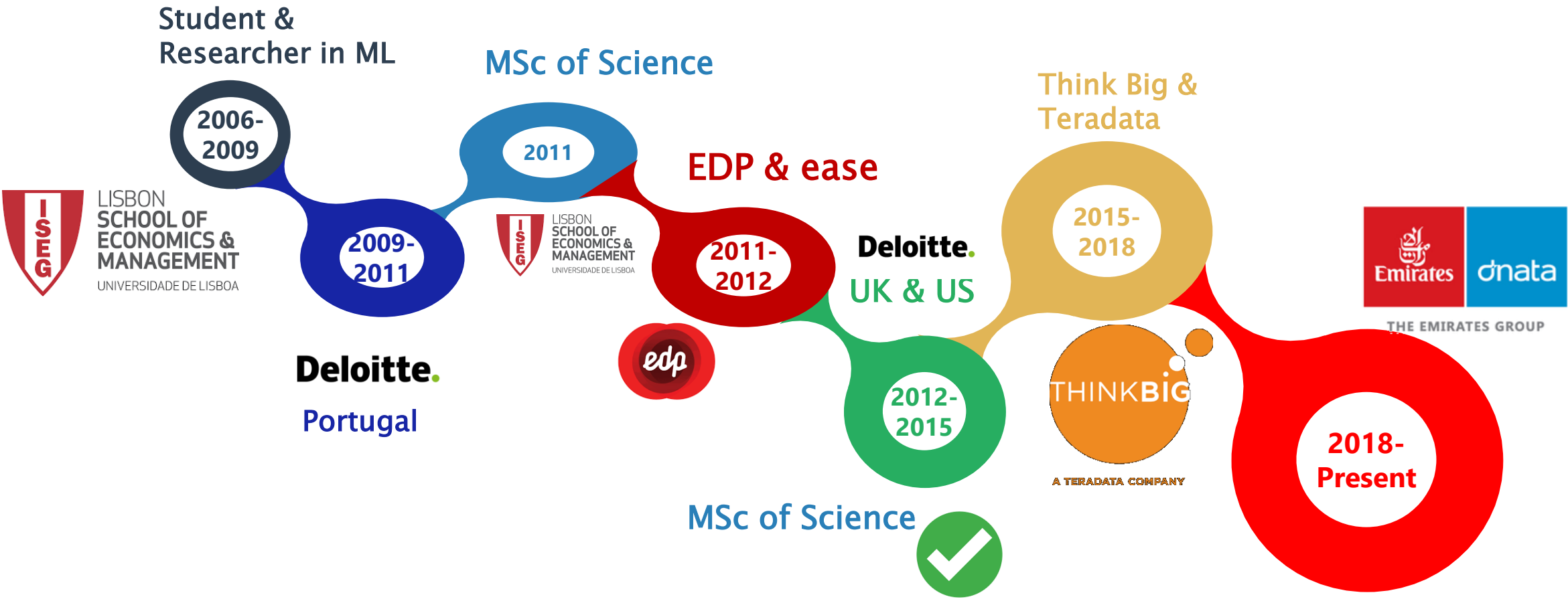


# DS Day 2019

Scaling Data Science in a “traditional” Enterprise  
Eliano Marques, VP Data Science, The Emirates Group

- 
- A vertical dashed line with six circles, each corresponding to an agenda item.
- **My 2 min Journey to Date**
  - **What I've learned about Enterprises & DS**
  - **Foundation to Scale & Deliver**
  - **Demand Management Model**
  - **Platform & DevOps Details**
  - **Real Use-Case**

# My 2 min Journey to Date



# The maturity levels (A<sup>3</sup>) of Data Science across the Enterprise

R&D  
DS Team

Applied  
DS Team

A<sup>3</sup>utonomous

- All Business & Digital applications speak "Data Science"
- Execs and Business Leaders driving the agenda of "Data Science"
- Data Science Investment/ Revenue ratio no longer have 10 zeros before a number



A<sup>2</sup>mbitious

- Focus on foundation
- Cares less about sexy words and more about production
- Very Business and Processes oriented



A<sup>1</sup>spirational

- Hire 1 (or a few) Data Scientist and thinks problems are solved
- Does PoC (internal or external) and believes job is done
- Seems to be taken off but runs in circles, churn starts to appear



# Particularities of “traditional” Enterprises

“Slow”

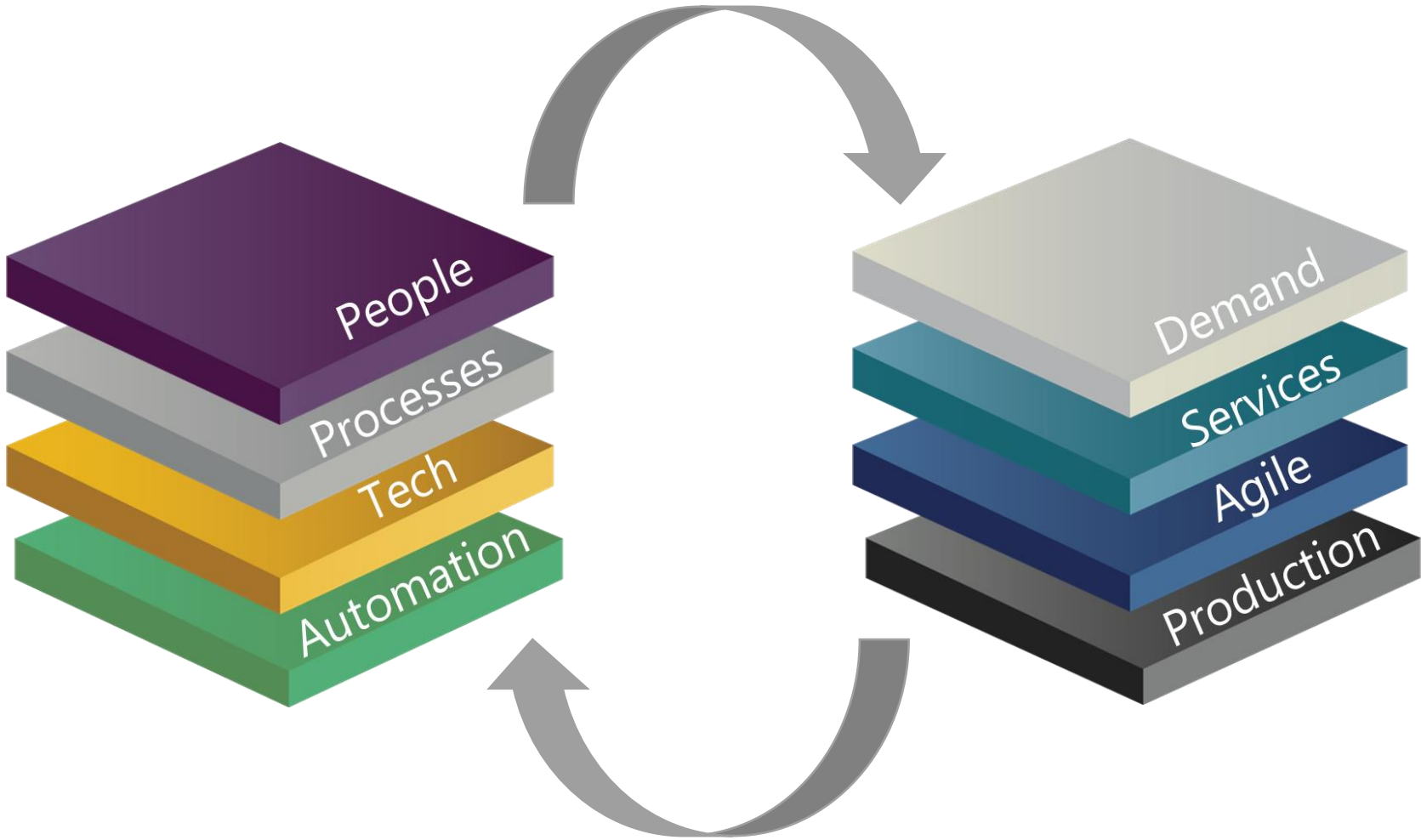
“Politically  
complex”

“Never ready  
to change”

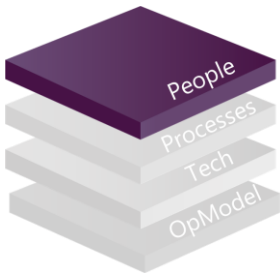


Foundation

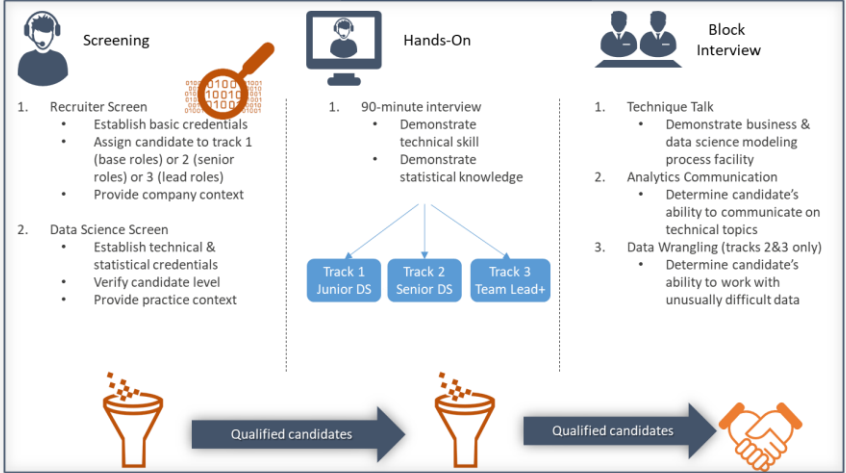
Delivery



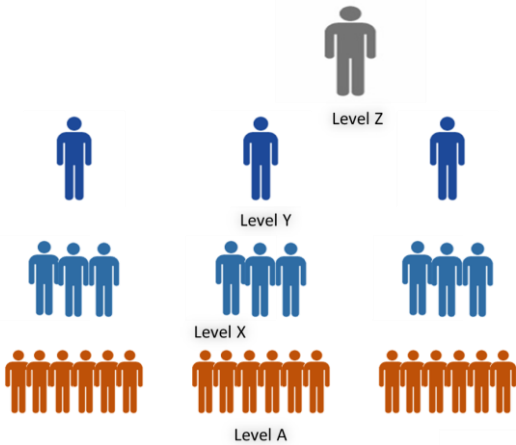
# Building the foundation - People



Recruit well  
& with  
standards

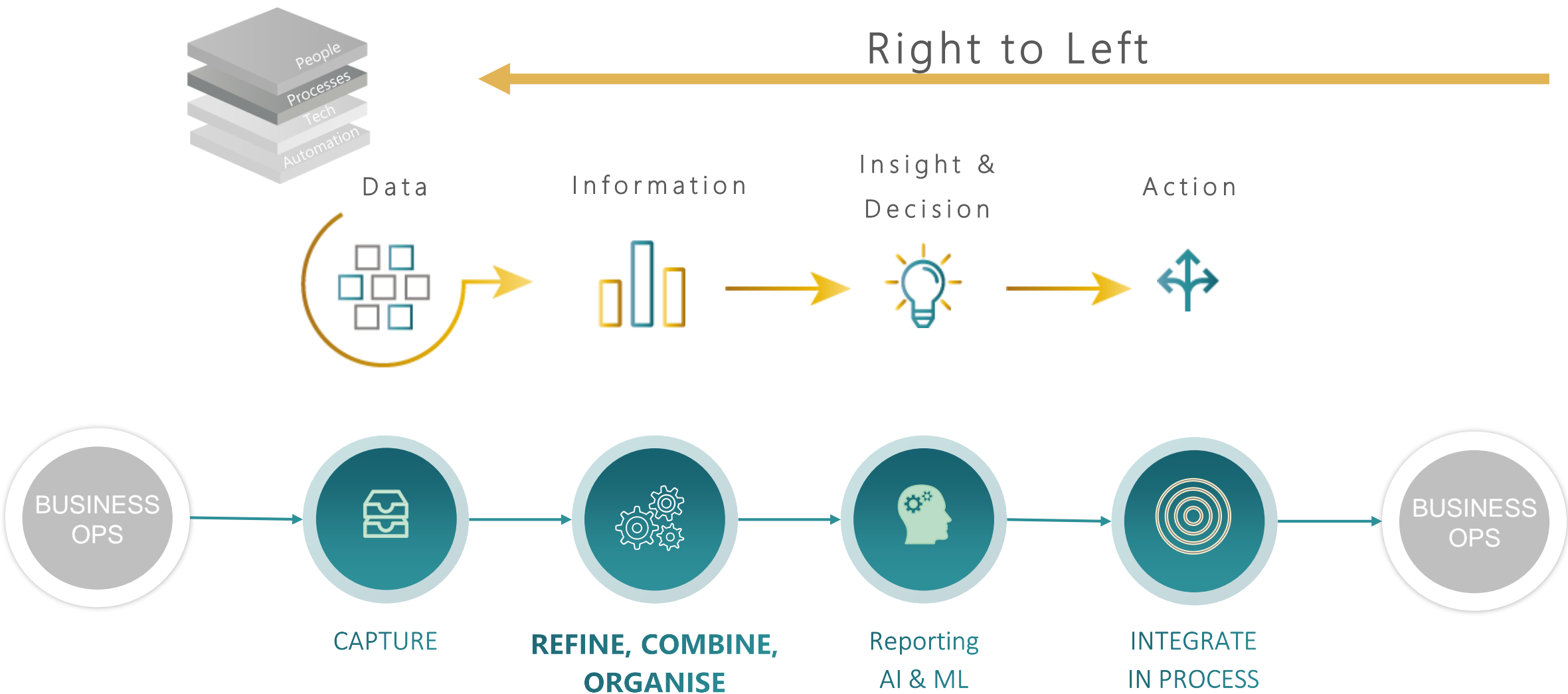


Organise to  
grow and win



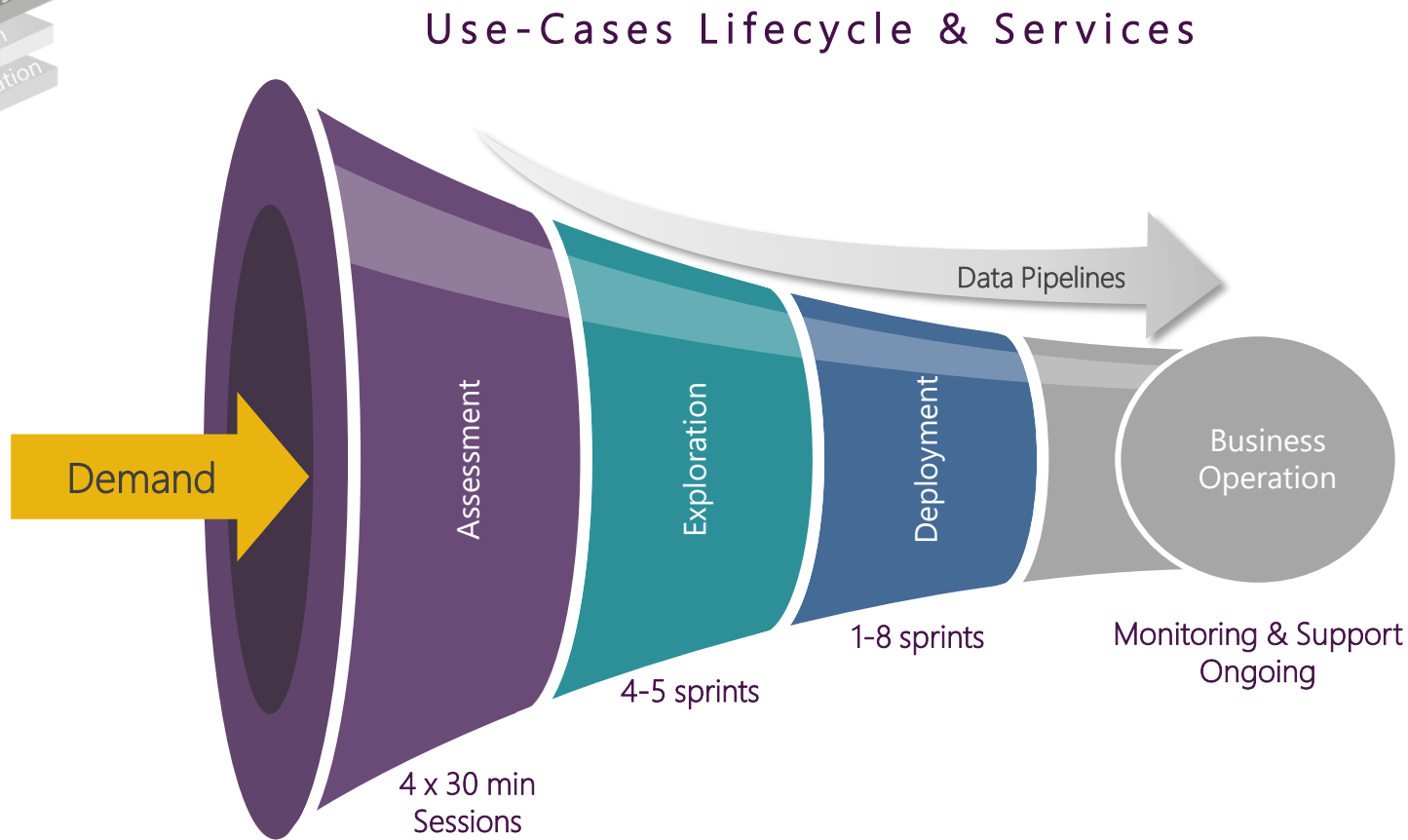
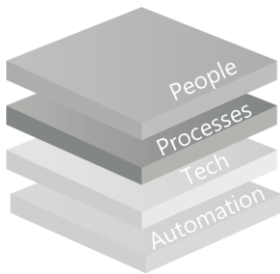
Invest in  
training paths

	Consulting Track	Data Science Technology Track	Data Science Track	Emirates Group Domains & Business Units
Level 1 - DS (Foundation)	DS110	DS120	DS130	DS140
Assessment 1	A-DS110	A-DS120	A-DS130	A-DS140
Level 2 - SDS (Intermediate)	DS210	DS220	DS230	DS240
Assessment 2	A-DS210	A-DS220	A-DS230	A-DS240
Level 3 - Team Lead (Advanced)	DS310	DS320	DS330	DS340
Assessment 3	A-DS310	A-DS320	A-DS330	A-DS340
Level 4 - VP+ (Expert)	DS410	DS420	DS430	DS440
Assessment 4	A-DS410	A-DS420	A-DS430	A-DS440



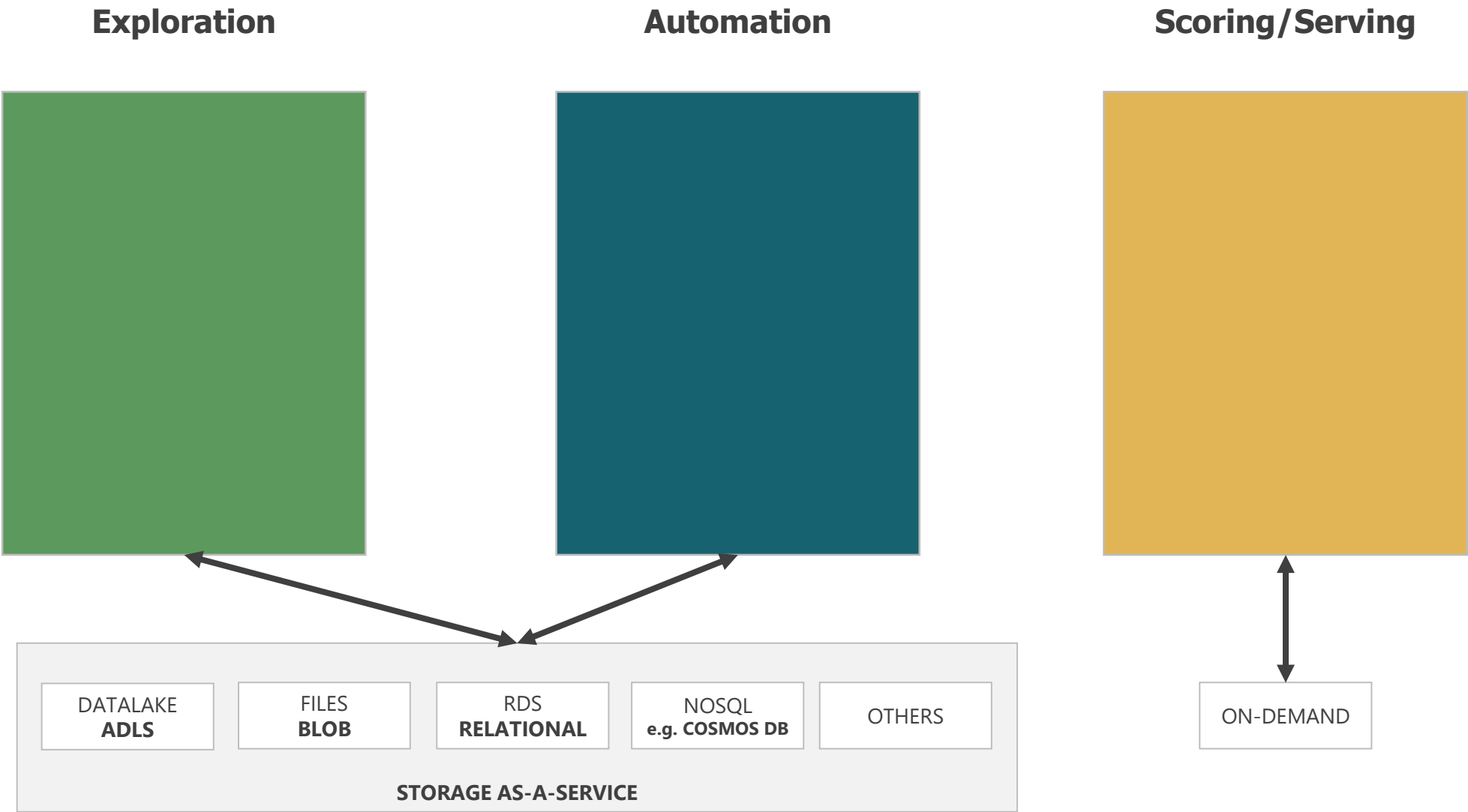
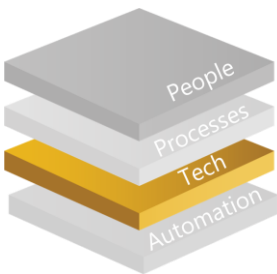


# Delivering value fast in production - Demand

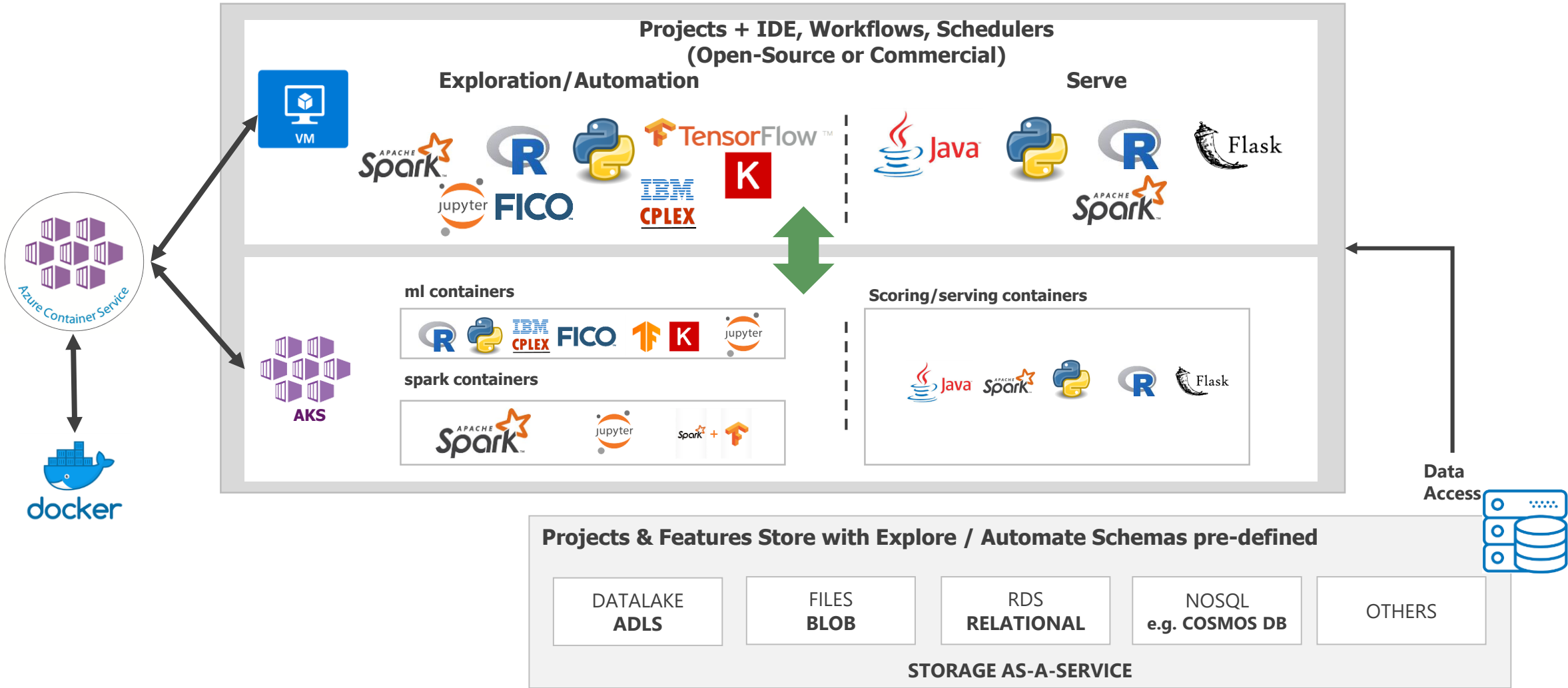


# Building the foundation – Tech

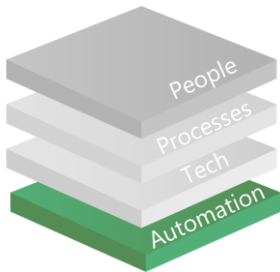
## High-level architecture



Bridge Concept

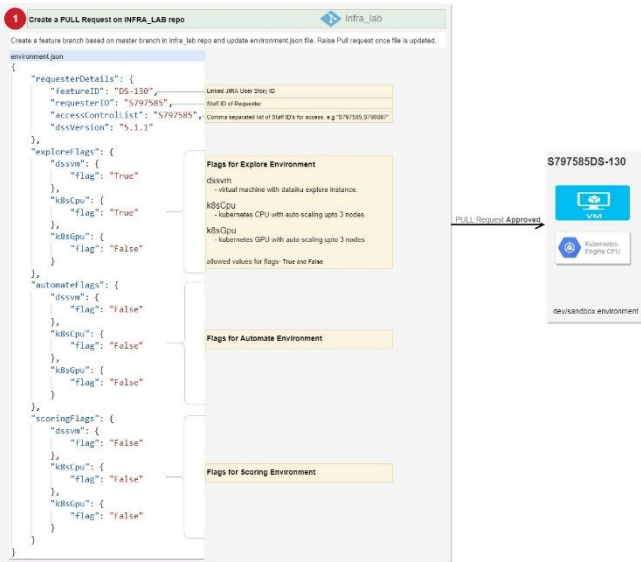


# Building the foundation - Automation

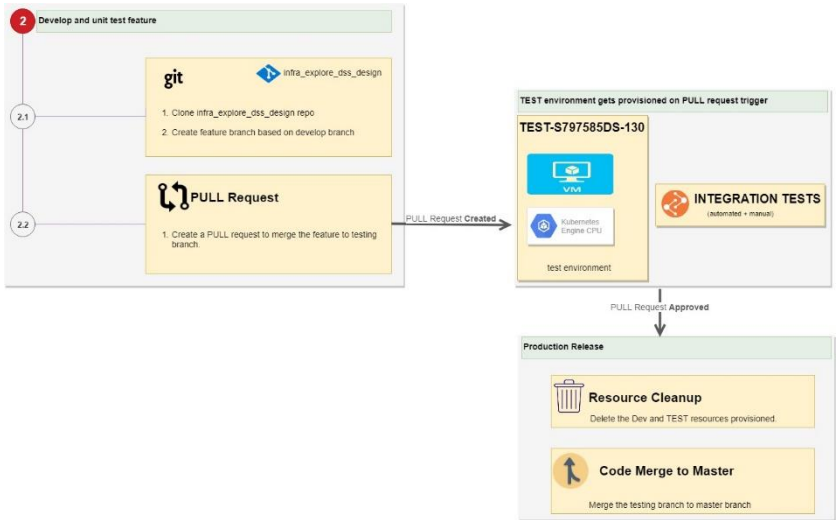


Platform

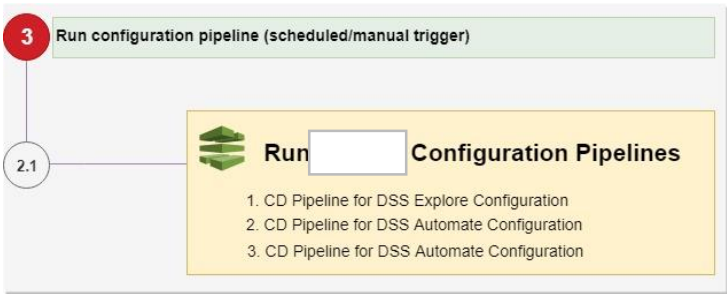
## Dev Automation



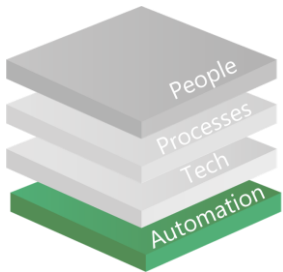
## Test Automation



## Prod Automation



# Building the foundation - Automation



Platform

1

Create a PULL Request on INFRA\_LAB repo

infra\_lab

Create a feature branch based on master branch in infra\_lab repo and update environment.json file. Raise Pull request once file is updated.

environment.json

```
{
  "requesterDetails": {
    "featureID": "DS-130",
    "requesterID": "S797585",
    "accessControllist": "S797585",
    "dssVersion": "5.1.1"
  },
  "exploreFlags": {
    "dssvm": {
      "flag": "True"
    },
    "k8sCpu": {
      "flag": "True"
    },
    "k8sGpu": {
      "flag": "False"
    }
  },
  "automateFlags": {
    "dssvm": {
      "flag": "False"
    },
    "k8sCpu": {
      "flag": "False"
    },
    "k8sGpu": {
      "flag": "False"
    }
  },
  "scoringFlags": {
    "dssvm": {
      "flag": "False"
    },
    "k8sCpu": {
      "flag": "False"
    },
    "k8sGpu": {
      "flag": "False"
    }
  }
}
```

Linked JIRA User Story ID

DS-130

Staff ID of Requester

S797585

Comma separated list of Staff ID's for access. e.g "S797585,S760667"

S797585

Flags for Explore Environment

dssvm  
- virtual machine with dataiku explore instance.

k8sCpu  
- kubernetes CPU with auto scaling upto 3 nodes.

k8sGpu  
- kubernetes GPU with auto scaling upto 3 nodes.

allowed values for flags- True and False.

Flags for Automate Environment

Flags for Scoring Environment

PULL Request Approved

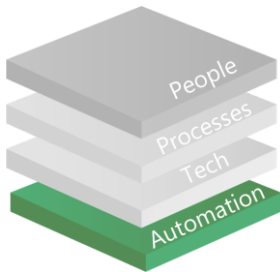
S797585DS-130

VM

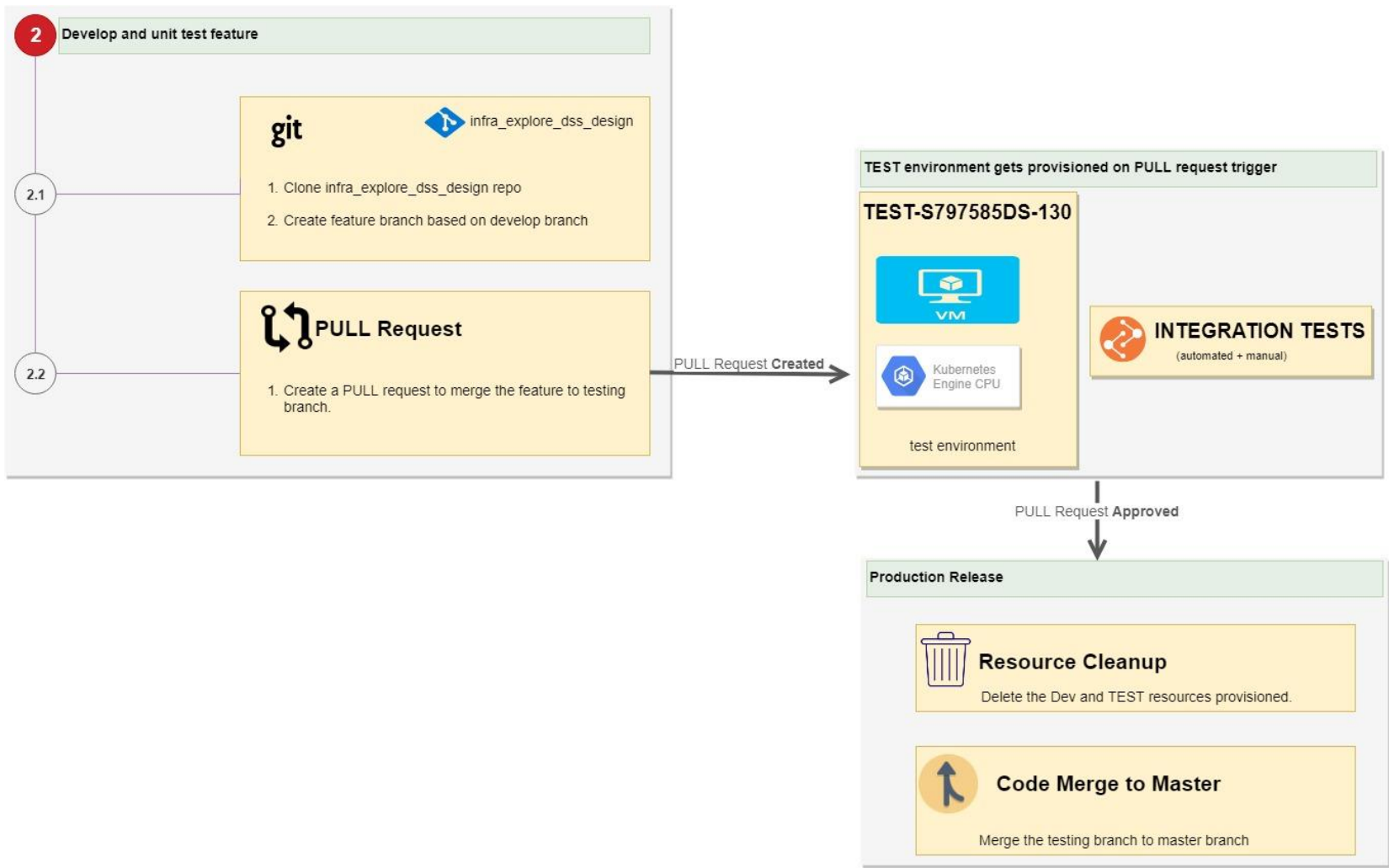
Kubernetes Engine CPU

dev/sandbox environment

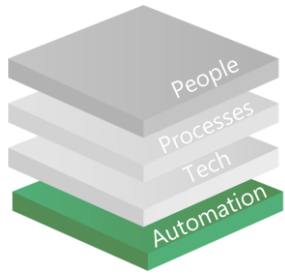
# Building the foundation - Automation



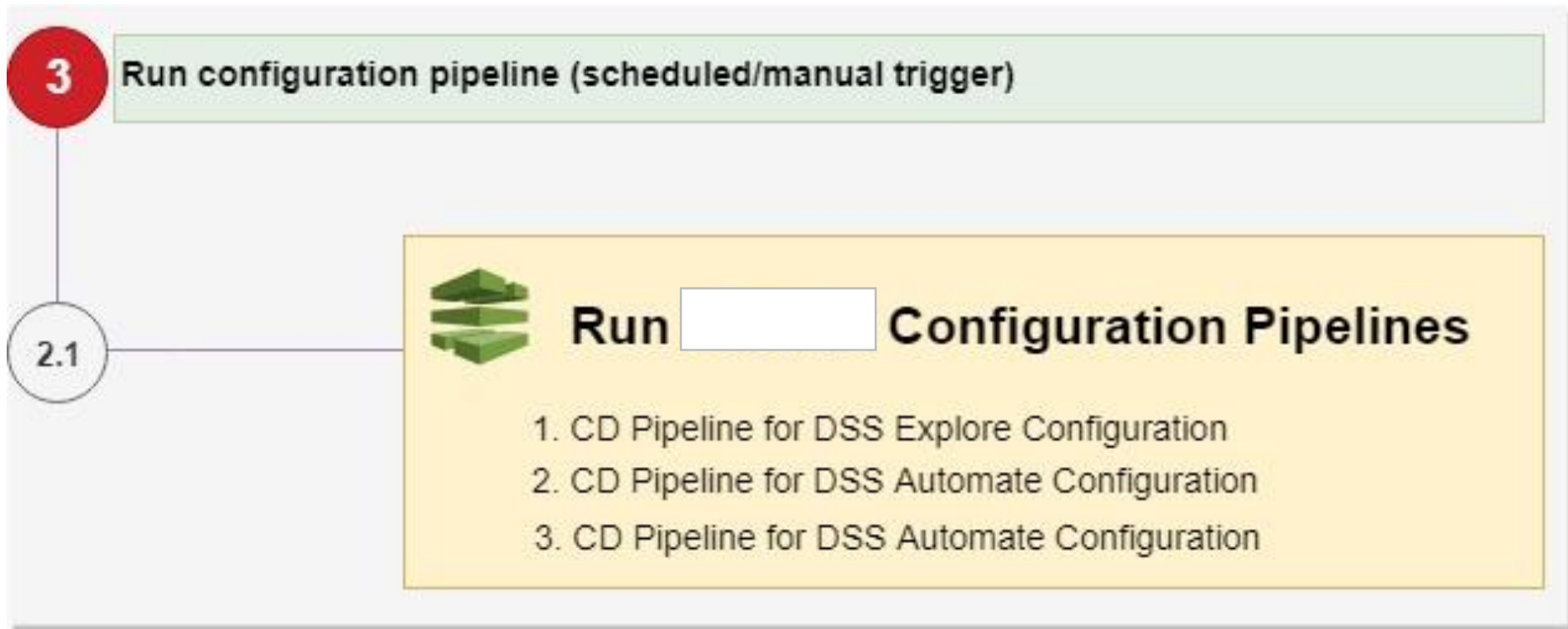
Platform





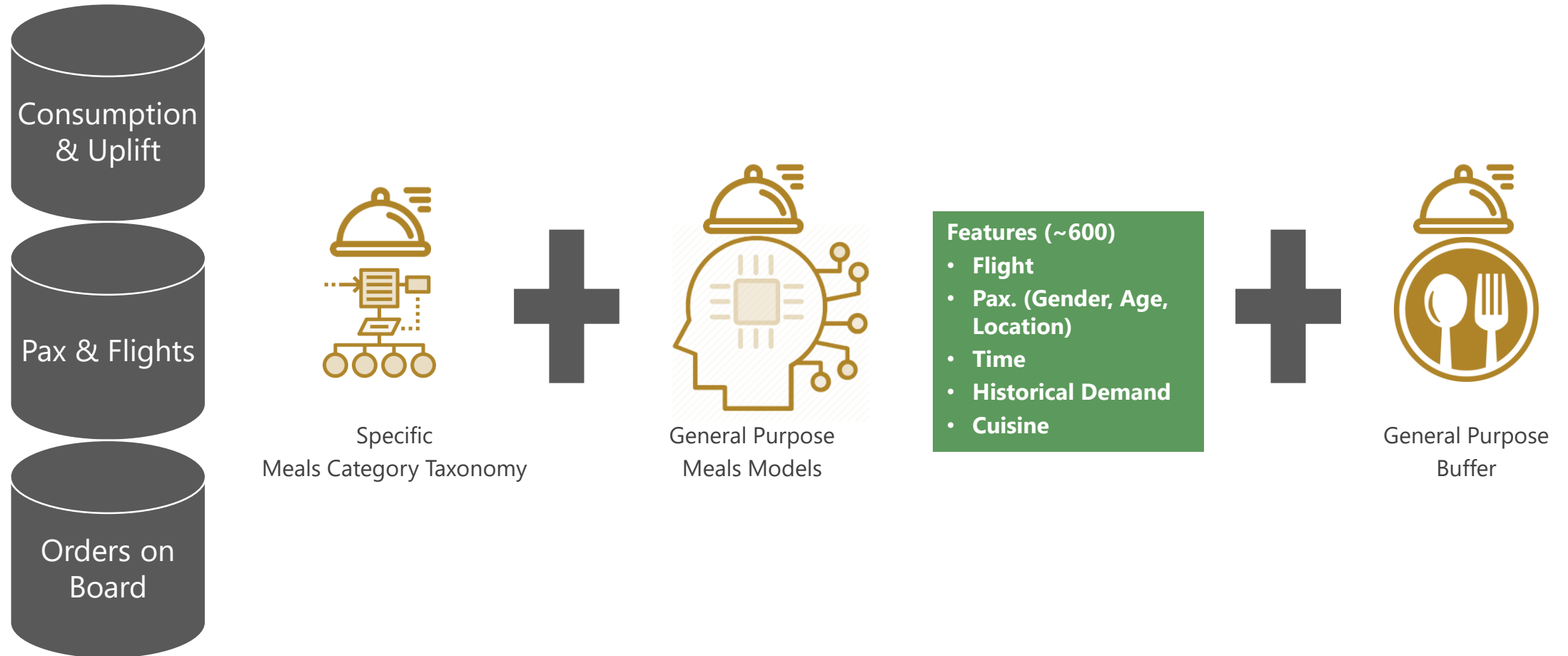


Platform

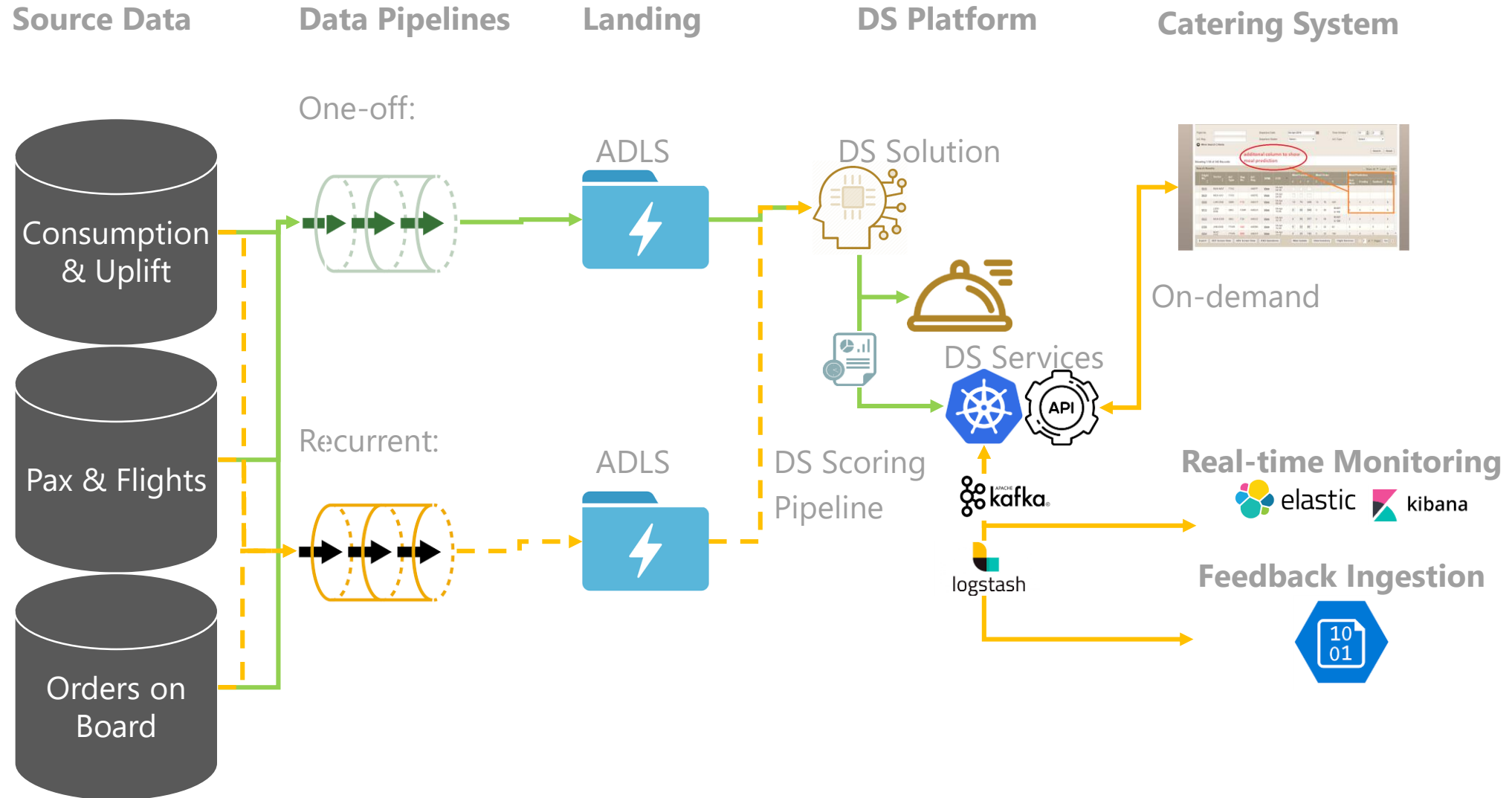


# Real use-case

## Data Science Approach



# Real use-case







# Thank you

I hope you had fun, if not, use this:  $\textit{fun} = 2 \cos \frac{1}{2}(\alpha + \beta) \cos \frac{1}{2}(\alpha - \beta)$   
and try again!