

DataOps 101

A Better Way to Develop and Deliver Data Analytics

June 20th, 2024

About Me

Experience

- 11+ years with Microsoft BI stack
- 8+ years with PowerPivot/Power BI
- 5+ years with Azure DevOps

Credentials

- MCSE: Productivity
- MS: PBI Data Analyst
- PMP



Contact Me



@JKerski



John Kerski



Blog

Slides:



Agenda

- Background
- What is DataOps
- Why it Matters
- How it can Help You

Scenario

- Alex – The Boss



- Bob – The Data Analyst



Day 1

Alex

Great job @Bob, the new Power BI Report is going over well with the executives!



Bob

Thanks! That's great to hear!

Day 2

Alex **IMPORTANT!**

@Bob, The executives are in a meeting and the slicer doesn't have all the options like it did yesterday, can you fix it ASAP?

Bob
I'm on it!



Alex

Thanks for saving the day superhero

Day 3

Alex **IMPORTANT!**

@Bob, Charlie in Accounting looked at the new Power BI Report and he says the numbers in the table visual don't add up correctly when you update that slicer you fixed yesterday?

Bob

I'm on it... may take two days because I need to test.

Alex

Let's get that fixed today.



Day 4

Alex **IMPORTANT!**

I got a call from an executive, the KPI card that uses the same measure in the table you fixed yesterday looks wrong.

Bob
I'm on it 😞



Alex
Let's not have more errors today.

Day 5

Alex

Accounting called again, the model has a new column in it, did you do that?

Bob

No, let me check.

Alex

Alright... rough week



Show of Hands

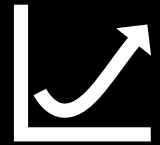


Risks

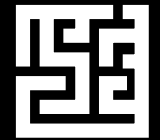


Issues

Data is a Strategic Asset but...



Demand is High



Complex Data Lineage



Data Quality Issues



Defects & Errors



Slow Delivery



More Cost



Disappointment

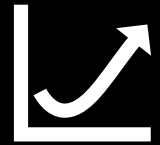
DataOps

What is DataOps?

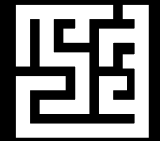


Why It Matters

Data is a Strategic Asset but...



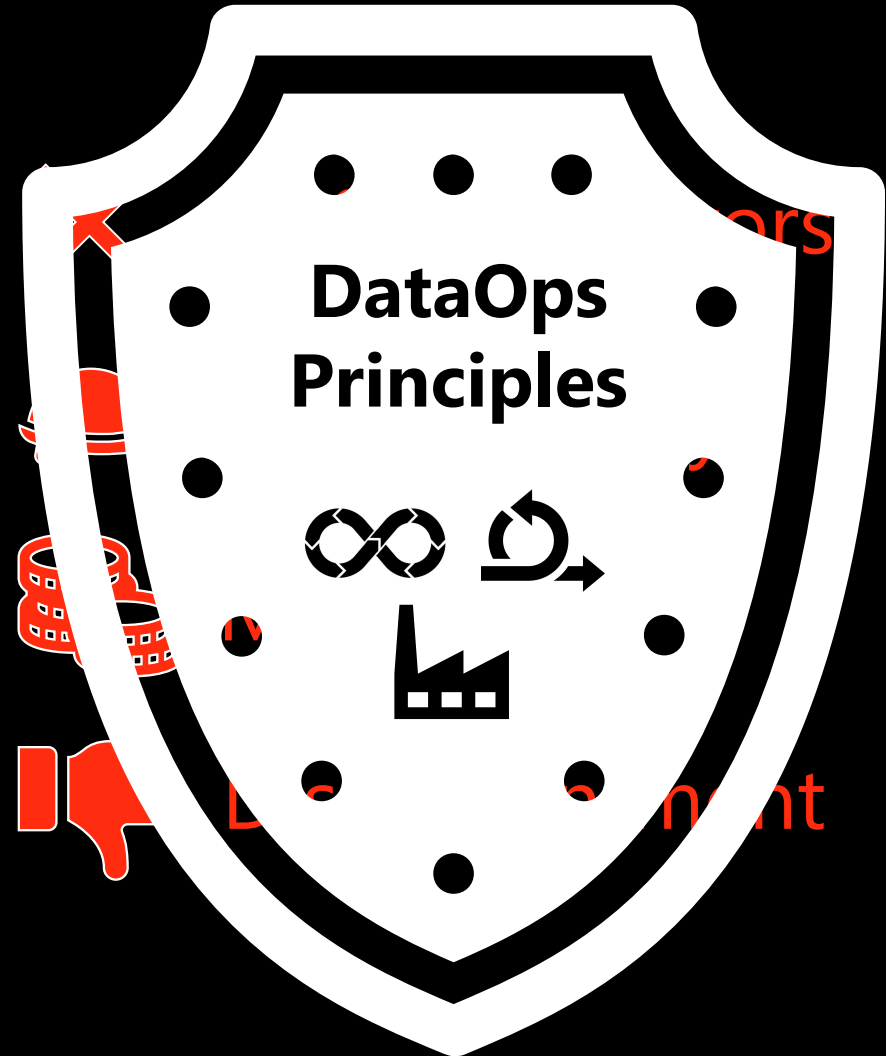
Demand is High



Complex Data Lineage



Data Quality Issues



How can it help me?

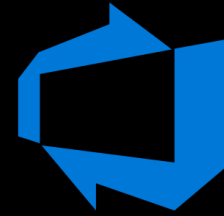
Make it Reproducible

Reproducible results are required and therefore **we** **version** everything: data, low-level hardware and software configurations, and the code and configuration specific to each tool in the toolchain.

Learn Git

Get your work into Version Control

Version Control Path



PBIP/Git Integration

The Future



pbi-tools

Next Step to Dev Mode




Git-LFS


First Step to Dev Mode





OneDrive/SharePoint


Low Barrier to Entry


 Executive Dashboard - real real.pbix


 Executive Dashboard (28).pbix


 Executive Dashboard (26).pbix


 Executive Dashboard (25).pbix


 Executive Dashboard (24).pbix


 Executive Dashboard (23).pbix


 Executive Dashboard (22).pbix


 Executive Dashboard (21).pbix


 Executive Dashboard (20).pbix


 Executive Dashboard (19).pbix


 Executive Dashboard (18).pbix


 Executive Dashboard (17).pbix


 Executive Dashboard (16).pbix


 Executive Dashboard - real copy.pbix


 Executive Dashboard - save.pbix


 Executive Dashboard - real.pbix


 Executive Dashboard - (50).pbix


 Executive Dashboard - (49).pbix


 Executive Dashboard - (48).pbix


 Executive Dashboard - (47).pbix


 Executive Dashboard - (46).pbix


 Executive Dashboard (7).pbix


 Executive Dashboard (6).pbix


 Executive Dashboard (5).pbix


 Executive Dashboard (4).pbix


 Executive Dashboard (3).pbix


 Executive Dashboard (2).pbix


 Executive Dashboard (1).pbix


 Executive Dashboard.pbix


 Executive Dashboard (30).pbix


 Executive Dashboard (30) - Copy.pbix


 Executive Dashboard (30) - Copy (1).pbix


 Executive Dashboard (30) - backup.pbix


 Executive Dashboard (20) - backup.pbix


 Executive Dashboard (5) - proposal.pbix


 Executive Dashboard (5) - proposal final draft.pbix


 Executive Dashboard (6) - final final.pbix


 Executive Dashboard - My Edits.pbix


 Executive Dashboard - DAX fix.pbix

 Executive Dashboard - Prototype.pbix

 Executive Dashboard (29).pbix

 Executive Dashboard (31).pbix

 Executive Dashboard (32).pbix

 Executive Dashboard - maybe fix.pbix

Quality is Paramount

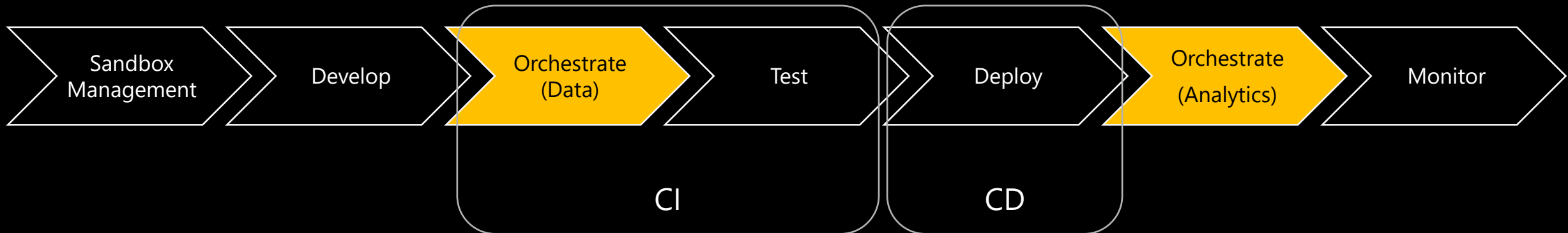
Analytic pipelines should be built with a foundation capable of **automated detection of abnormalities** and security issues in code, configuration, and data, and should provide continuous feedback to operators for error avoidance.

*When you Find an Error, Build a Test
Tests are your Safety Net*

Orchestrate

The beginning-to-end orchestration of data, tools, code, environments, and the analytic teams work is a key driver of analytic success.

Continuous Integration
Continuous Deployment
Scripts over Clicks



Monitor quality and performance

Our goal is to have performance, security and quality measures that are **monitored continuously** to detect unexpected variation and generate operational statistics.

Catch the Errors Before Your Customers Do
Practice What You Build

Reduce Heroism

As the pace and breadth of need for analytic insights ever increases, we believe analytic teams should strive to reduce heroism and create sustainable and scalable data analytic teams and processes.

Avoid Burnout

Recap – How Can it Help You

1. **Make It Reproducible** - Learn/Embrace Git
2. **Quality is Paramount** - Find an Error, Build a Test
3. **Orchestration** - Implement CI/CD
4. **Monitor Quality and Performance** - Catch Errors Before Your Clients Do
5. **Reduce Heroism** – Avoid Burnout

DataOps Aphorisms

- It is not about the fear of making a change; it is about the removal of fear from making changes.
- It is not about data quality; it is about low rates of error.
- It is not about the next analysis; it is about building a system that makes it easier to answer the next analysis.
- You spend all day helping your customers be more analytic, so why are you so un-analytic about how you run your data analytics teams?
- Don't solve problems, figure out how to never create them in the first place.

Resources

- General
 - [DataOps Cookbook](#)
 - [DataOps – Chris Bergh](#)
- Version Control
 - [GitHub Desktop Tutorial](#)
 - [YouTube - Version Control](#)
 - [pbi-tools](#)
- Testing & Orchestration
 - [YouTube - Testing](#)
 - [Visual Testing](#)
- Monitoring
 - [Template](#)

Questions?

