Praveen Bagare

AI & ML in TDM

PACIFIC NW SOFTWARE QUALITY CONFERENCE

THE FUTURE IS NOW

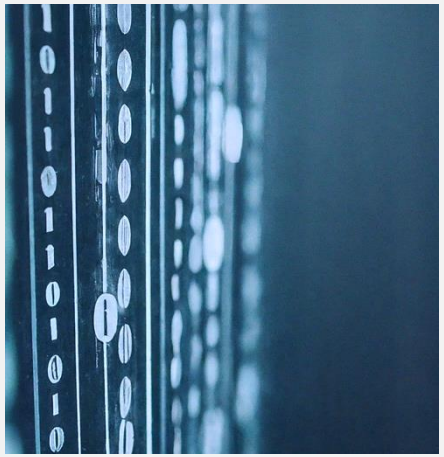
PNSQC.ORG

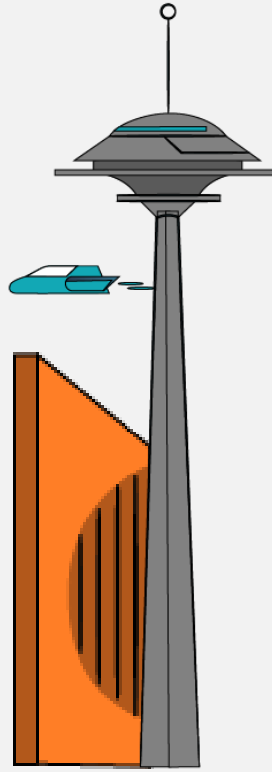
OCTOBER 14-16 2024

What is Test Data Management?

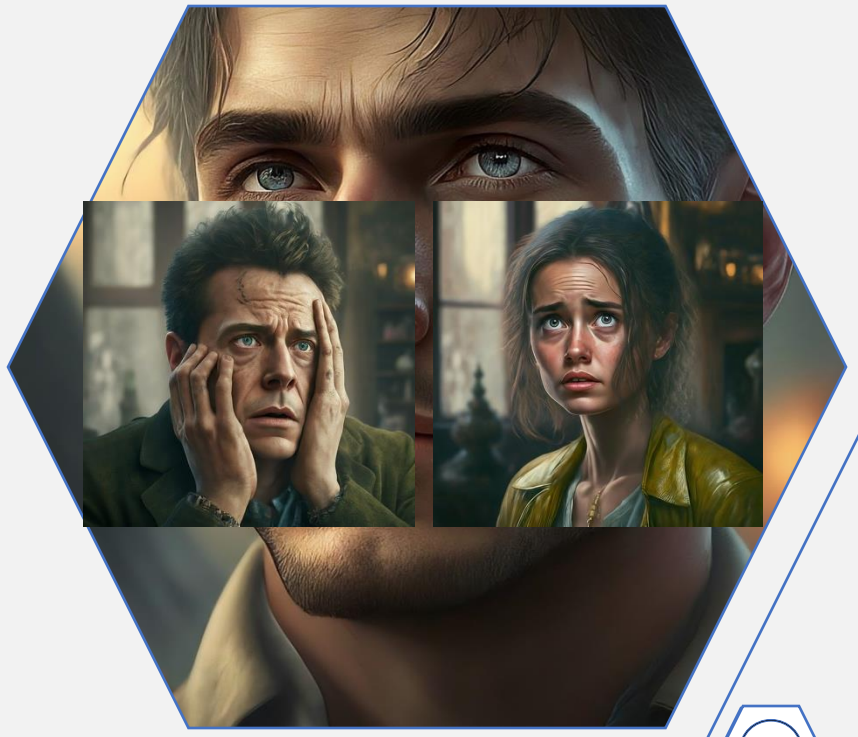
There are many definitions, here is one.

The process of provisioning data for testing ensuring it is of high quality, provisioned in suitable quantity, right format, appropriate environment, and in the stipulated time.

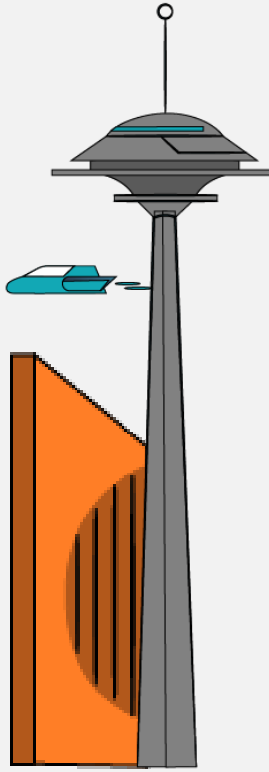




The image above was created using Imagen text-to-image model.

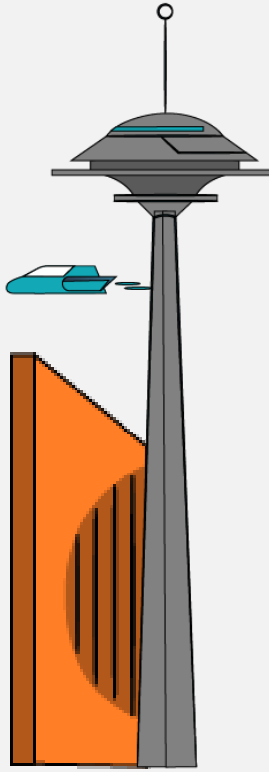


About 50% of a tester's effort is spent around sourcing, maintaining and managing test data.



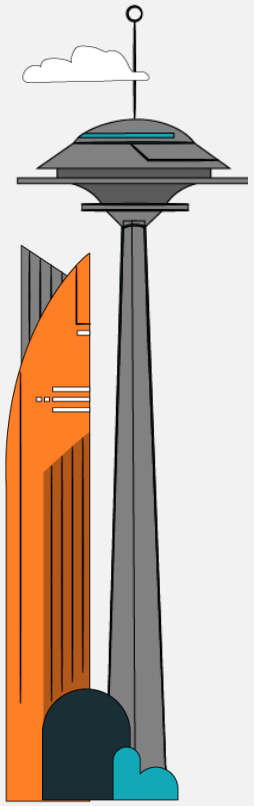
The image above was created using Imagen text-to-image model.

How do we solve this problem?



The image above was created using Imagen text-to-image model.

TDM Approach 1



Leverage Masking

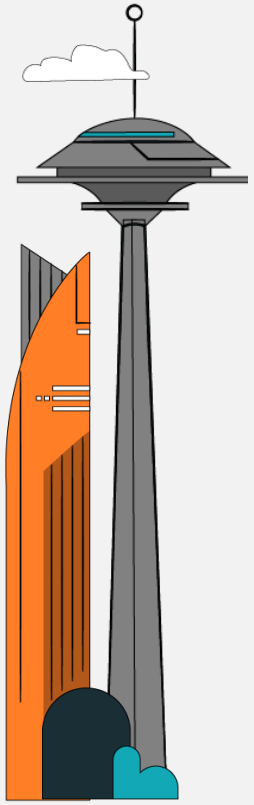
- Discover the data landscape
- Identify a tool
- Find Sensitive data
- Get stakeholder sign off
- Mask and load the data
- Get feedback & improve

Praveen Bagare

AI & ML in TDM

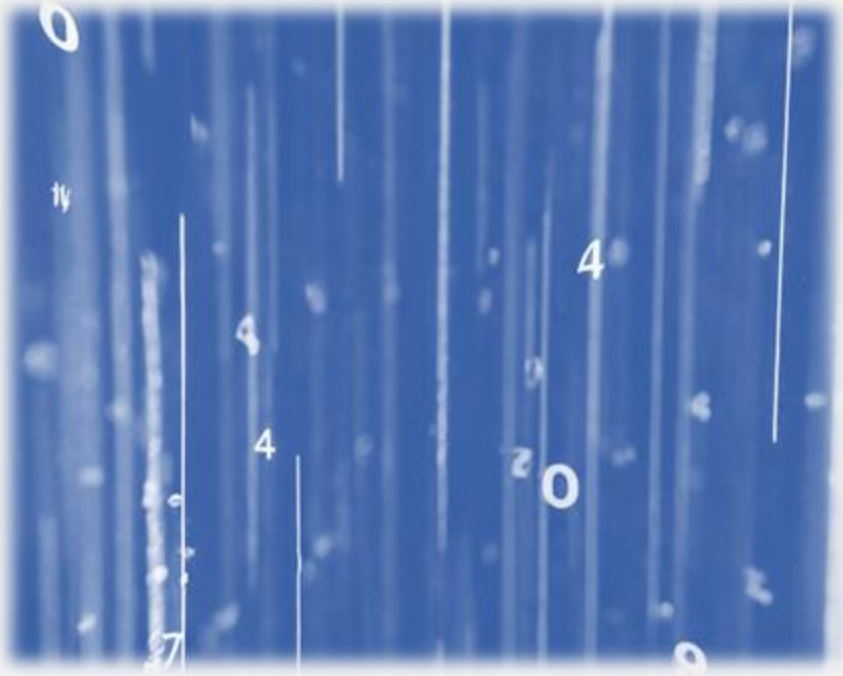
The image above was created using Imagen text-to-image model.

TDM Approach 2



Leverage Synthetic data generation

- Discovery - Understand metadata, relationships and data flow
- Identify right data generation tool
- Build generation rules, business logic, lookup etc.
- Generate & load synthesized data
- Get feedback & improve

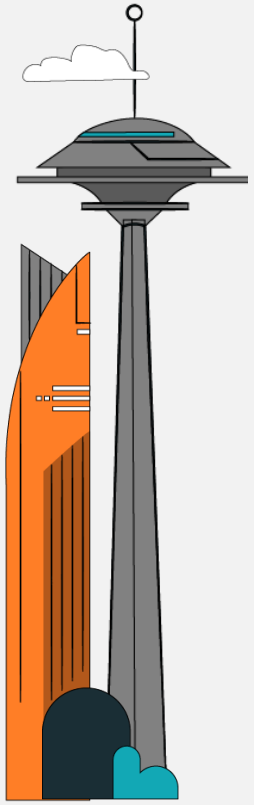


Praveen Bagare

AI & ML in TDM

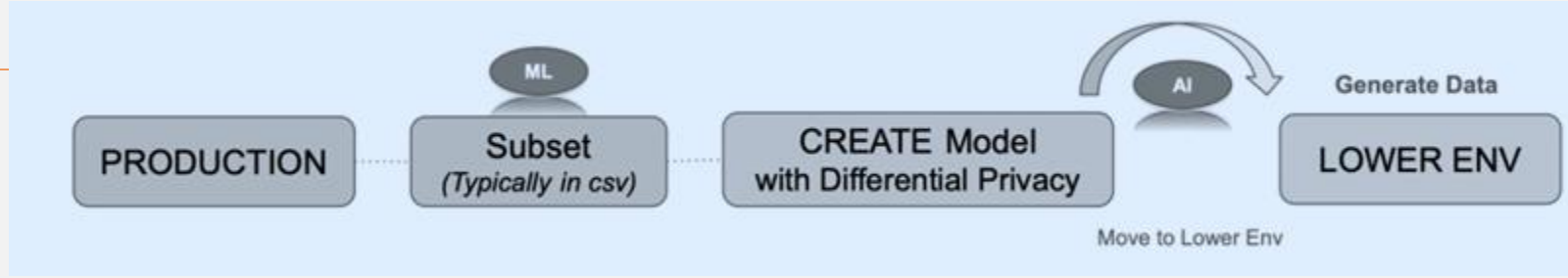
The image above was created using Imagen text-to-image model.

TDM Approach 3



Leveraging AI & ML data generation

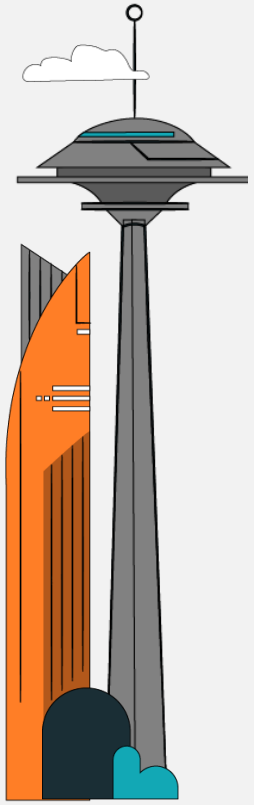
- Identify the data in scope *(subsetting)*
- Update config file *(number of epochs, relationships, pass through etc.)*
- Run machine learning *(Train)*
- Move the differential privacy applied model to desired environment.
- Run the synthetic data generation *(Infer)*



Praveen Bagare

AI & ML in TDM

The image above was created using Imagen text-to-image model.

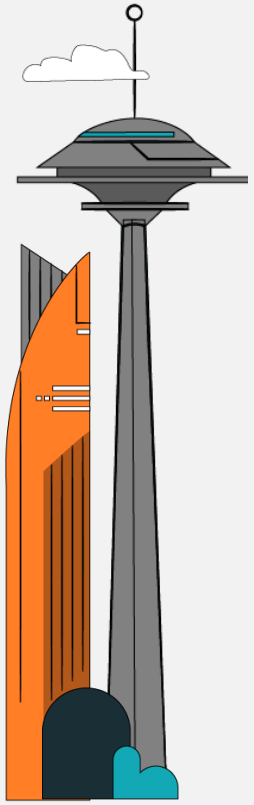


Approach 3 Demo

Leveraging AI and ML in data generation

Praveen Bagare

AI & ML in TDM



Please share your feedback!

Praveen Bagare

bagare@gmail.com

Praveen_Bagare@epam.com

https://www.linkedin.com/in/praveen-bagare-56b33169

Thank you!