Big Data's Influence on Product Design and Quality

Joseph Sefair

Agenda

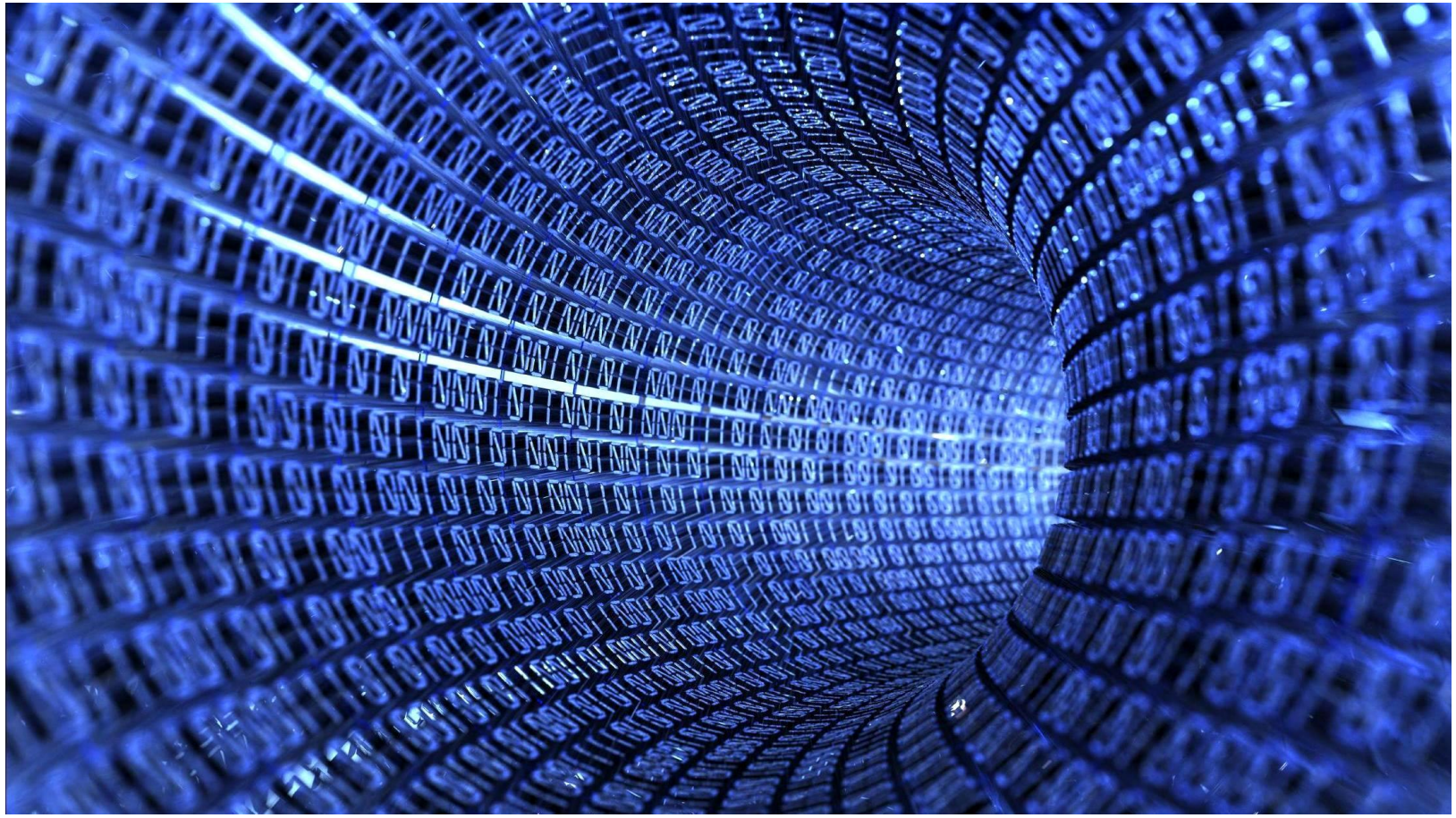
- **Big Data - Setting yourself up for success**
  - What are the basics – can you deal with the Big Data Chaos?
  - What problems are you really trying to solve?
  - Who (disciplines), what, and how
  - How programs evolve

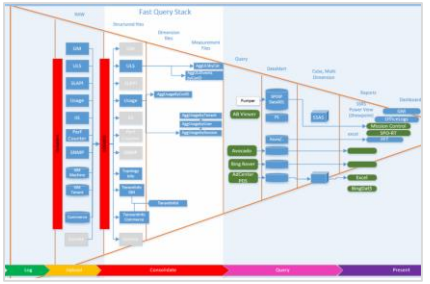
- **Case Studies / Examples of Evolution**
  - Microsoft Application ScoreCard Example
  - Office Anomaly Detection

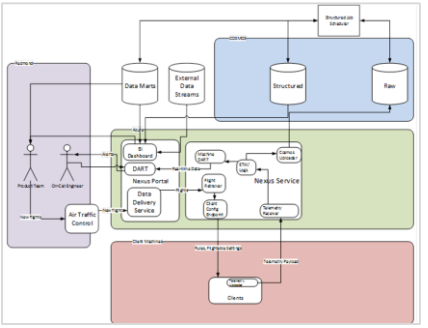
Big Data

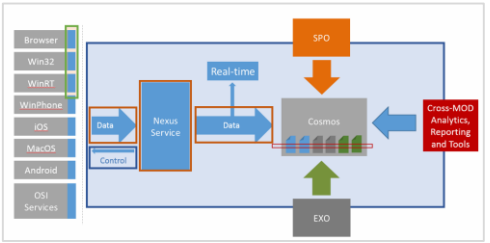
Setting yourself for success

Big Data






Data Marts
External Data Streams
Structured
Raw
Dashboard
DART
Nexus Portal
Data Delivery Service
Air Traffic Control
Nexus Service
Clients


Browser
Win32
WinRT
WinPhone
iOS
MacOS
Android
OSI Services
Data
Control
Nexus Service
Real-time
Data
SPO
Cosmos
EXO
Cross-MOD Analytics, Reporting and Tools

Can you organize your data?

What is the problem you are trying to solve?

What are the questions you are trying to answer?

Do you have the taxonomy that will help you ask these questions?

Does each discipline know their Data role?

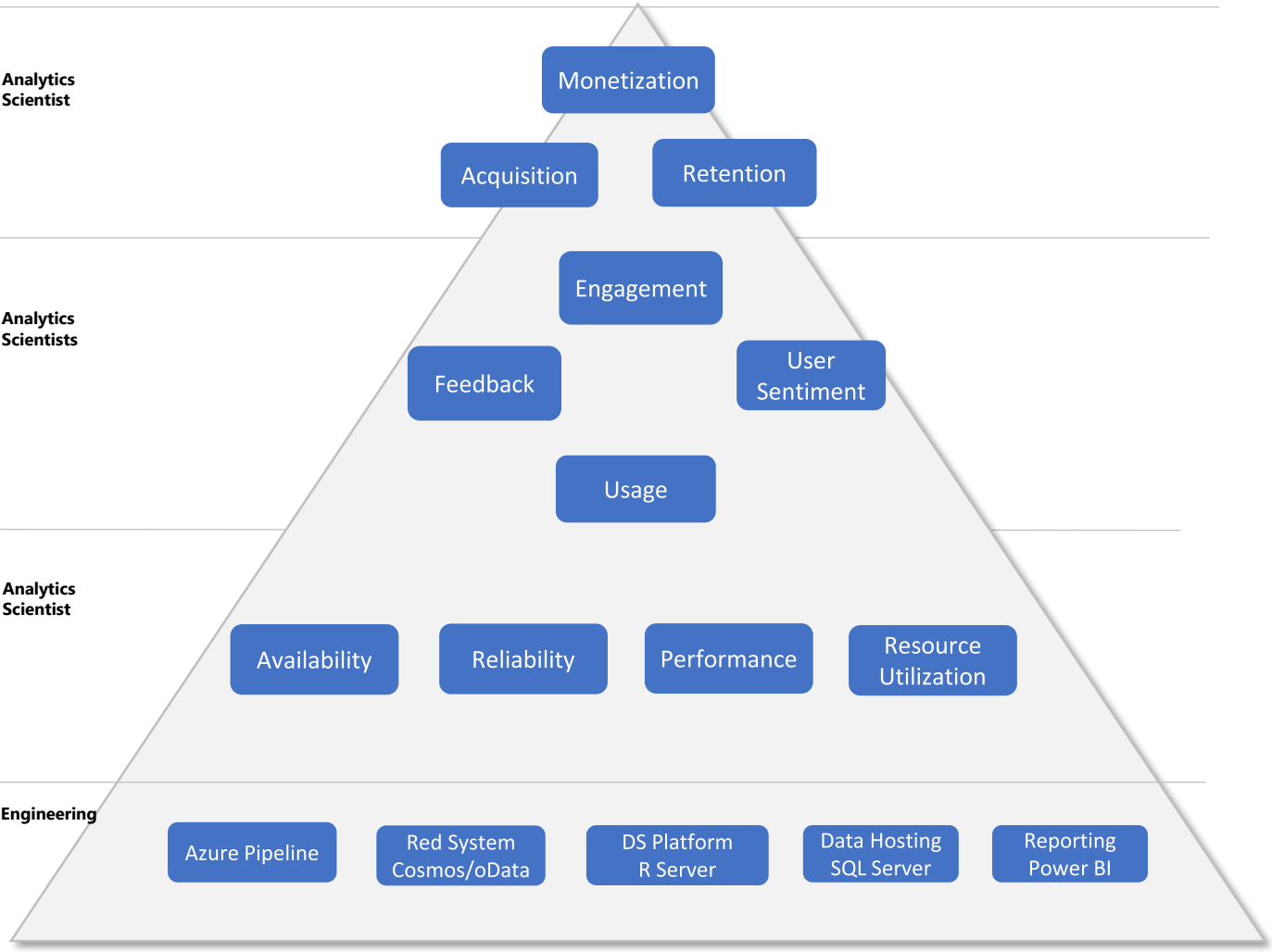
How can you improve

- Share knowledge across teams and orgs
- Taxonomy, semantics, common language
- De-Duplication of work across teams and orgs
- Collaboration across teams and orgs

Customer 'Delight'…experiences that span data and roles

	Measures	Discipline Partner
BUSINESS HEALTH How is the product /service contributing to Microsoft's bottom line and/or strategic goals.	**Attraction** Unaided awareness, Favorability, … **Engagement** New users, Unique users, Session duration, … **Retention** Repeat use, … **Monetization** Subscription conversion, Support costs…	**GPM** **Data Analytics** **Data Scientist**
Dependent upon…		
EXPERIENCE HEALTH How valuable is the product/service to customers – does it delight them *and* beat competitors.	**Overall Customer Delight and Satisfaction** Do users find our product delightful? Easy to use? Pleasurable? Feature usage – most, least, etc. Feedback, Sentiment, NSAT, NPS, … Subjective user ratings for top jobs Task completion rates, times	**PM** **Data Analytics** **Data Scientists**
Dependent upon…		
SYSTEM HEALTH What is the state of what we are building – are things working as designed and as compliance demands.	**Service up-time** **Feature Completeness** **Errors** **MTTR** **Performance** **Compliance**	**SWE** **Data Analytics** **Data Scientist**
Dependent upon…		
DATA ENGINEERING PLATFORM What is the platform we are using for the System – do we have the right data streams, reporting and experimentation platforms.		**Data Engineering**

Experiences



How programs evolve using data

HOW ARE WE DOING?

- KPIs defined, sources identified and first dashboard

WHY DID THIS HAPPEN?

- Refined metrics and self-service reporting

WHAT HAPPENS IF I DO THIS?

- Machine Learning and Predictive Analysis

PRODUCT HEALTH SPECTRUM

Contoso Application (Example): App edit / render 3D Objects

Customer 'Delight'…experiences measures example

Measures

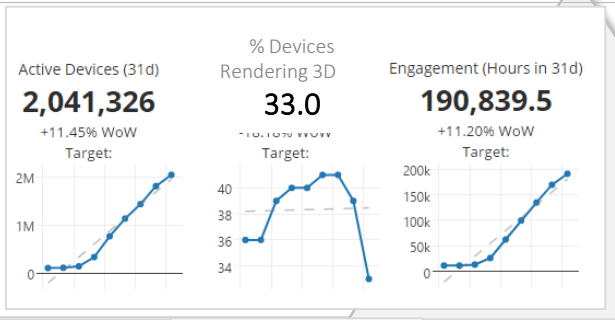
BUSINESS HEALTH

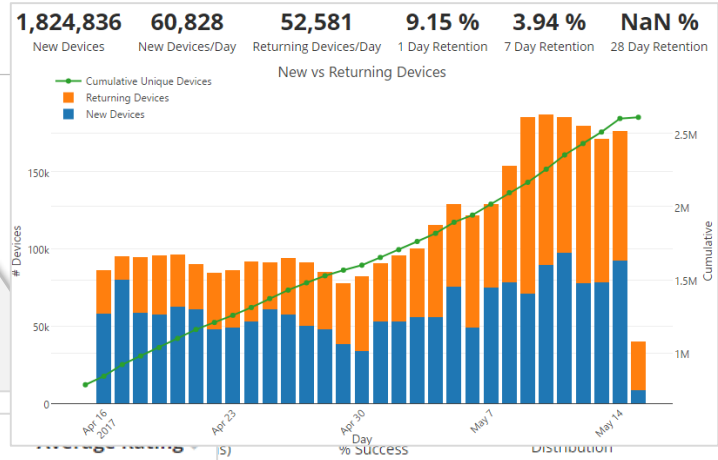
Dependent upon…

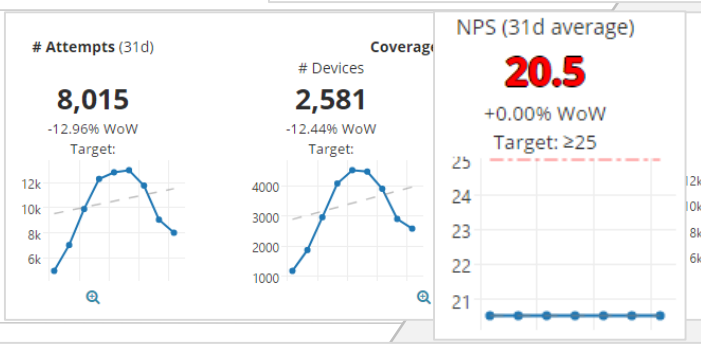
EXPERIENCE HEALTH

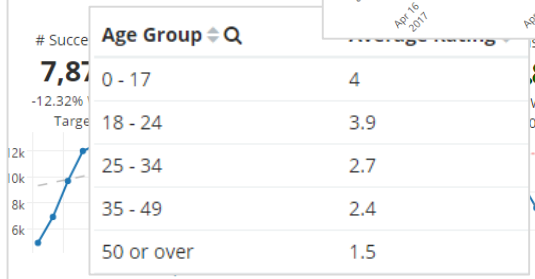
Dependent upon…

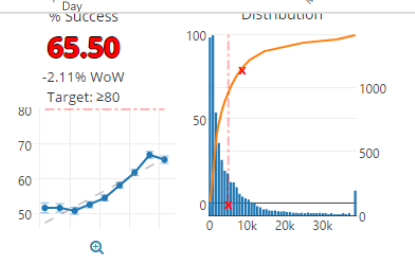
SYSTEM HEALTH

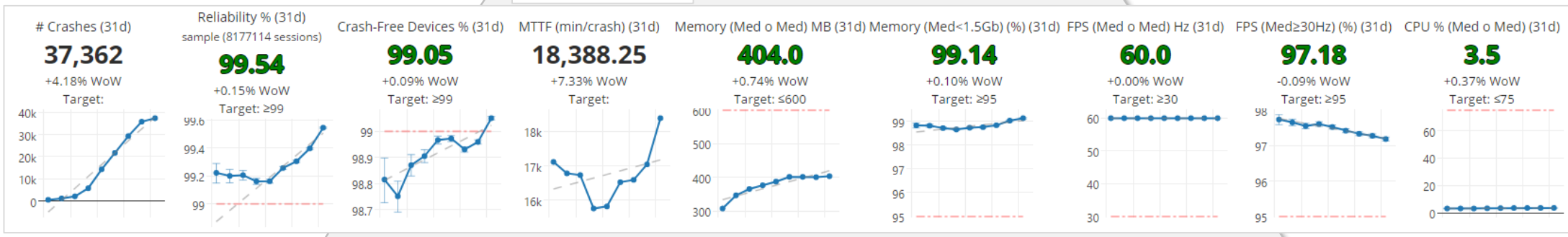












Customer 'Delight'…experiences measures example

Measures

BUSINESS HEALTH

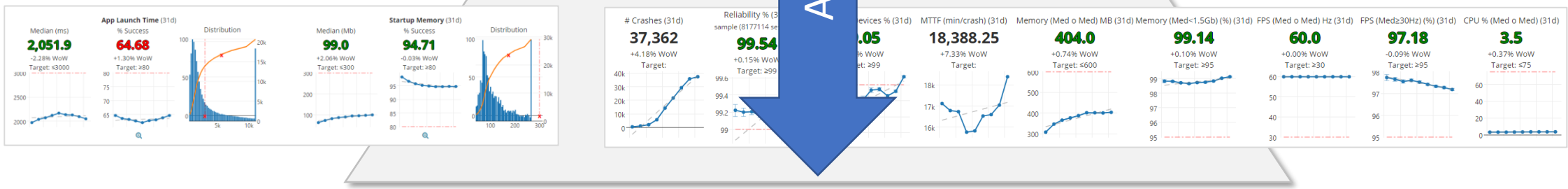
Dependent upon…

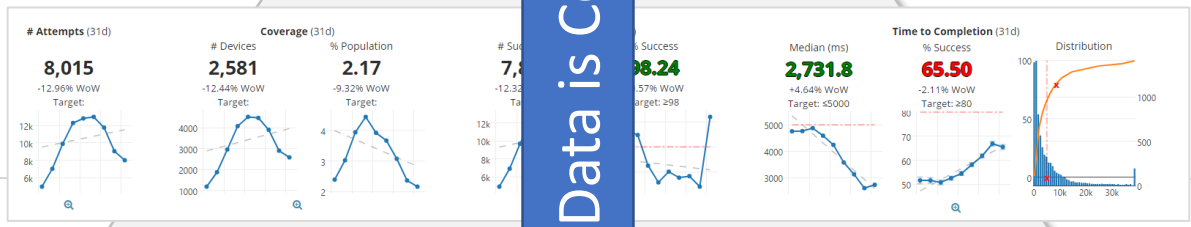
EXPERIENCE HEALTH

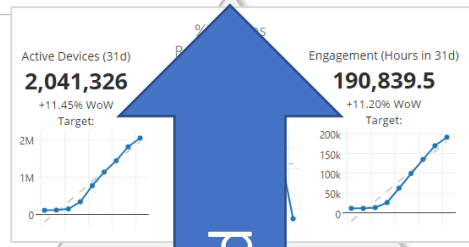
Dependent upon…

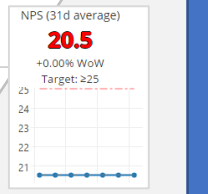
SYSTEM HEALTH

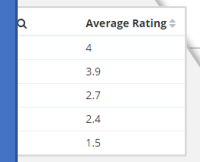
All Data is Correlated

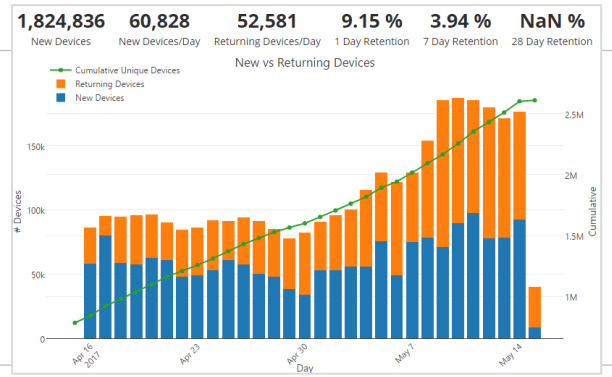












Anomaly detection

Office Reliability

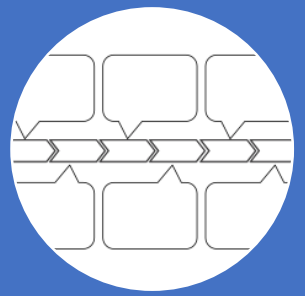
Anomaly Detection - Project Plan

Office Reliability - Example

How do we define ‘success’?

Detailed Walkthrough

- Activity reliability
- Activity performance
- Activity Volume
- Watson
- UAE



Timeline

Resources

What is ‘Anomaly detection (AD)’?

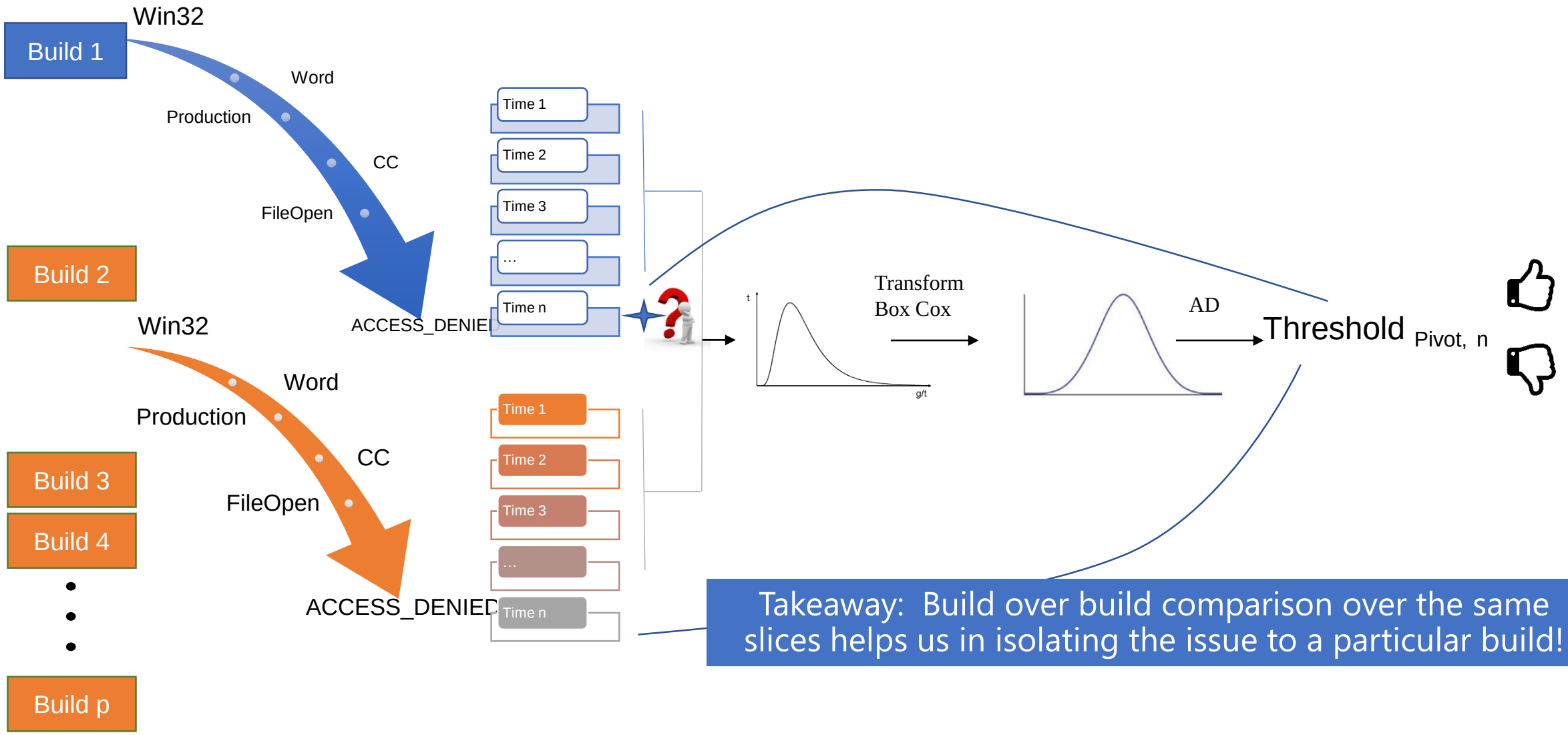
There are a lot of definitions out there –

“In data mining, **anomaly detection** (also **outlier detection**) is the identification of items, events or observations which do not conform to an expected pattern or other items in a dataset”

In our world,

- The model that represents a normal behavior from a given normal training set and testing the likelihood of a test instance to be generated from the learnt model
- Training involves learning behaviors from other builds and then test a given build against the dynamic thresholds

Massively Parallel Time Series AD Model



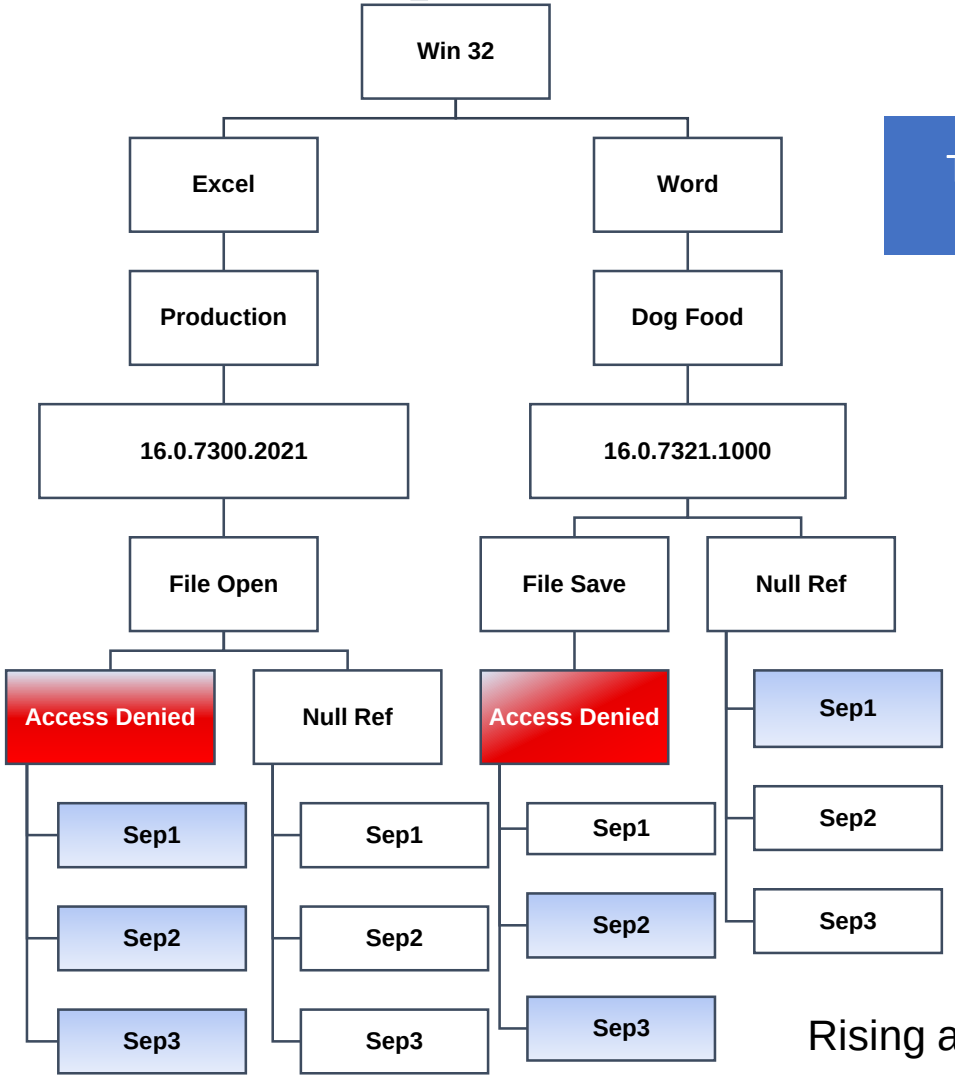
Takeaway: Build over build comparison over the same slices helps us in isolating the issue to a particular build!

Rolling up anomalous points to anomalous pivots

Takeaway: Grouping anomalous points to pivots makes it more actionable!

Anomalous points

Anomalous pivots



Single day spikes

All days anomalies

Rising anomalies

Our ‘success metrics’

- **Time To Detect** – This is a top level org metric, that measures our ability to detect ship-blocking issues earlier
  - $\sum_{k=0}^{n}$ Time when bug was created –
    Time when the issue was checked in (based on root cause)
- **Internal metrics**
  - Usage of the reports/offerings per team/GEM/GPM
    - # of users
    - # of views
  - # of Alerts created
  - # of Bugs created, sliced by ship-blocking v/s not

Thank You!

Questions?