

0

Measuring MLOps ROI & Team Productivity



Mikiko Bazeley
Head of MLOps
Featureform

Goals

- What are the North Star metrics of an effective ML and MLOps practice?
- How should MLOps and Data Science teams be organized to provide the most value to the organization?
- How do we connect the tactical, operational, & strategic layers of an MLOps practice?



"If you can't measure it, you can't manage it."



My BLUF

- Don't build an ML Platform unless:
 - You're post early-stage startup;
 - You have the ability to centralize;
 - You have a roadmap & strategy that involves adopting existing tools & you're willing to forgo building something shiny and new;
 - Your teams are willing to do enablement & adopt through influence.



My BLUF

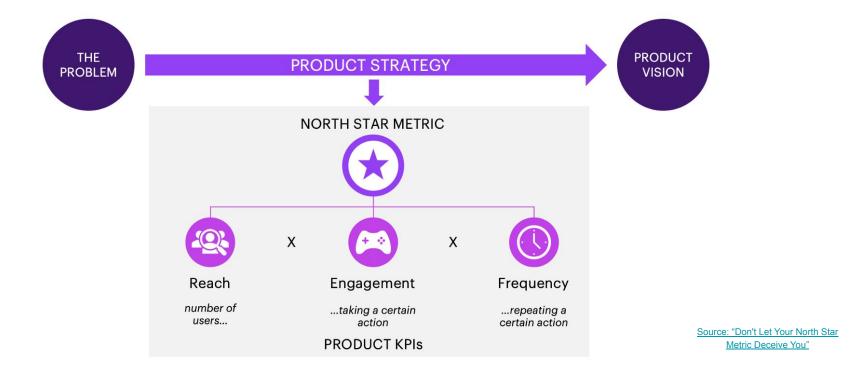
- If you do build a Platform, only adopt metrics that:
 - Your team can directly influence;
 - Are appropriate to the level of engagement;
 - Directly measure the behavior you're trying to capture.
- If you do decide to build an ML platform and an MLOps teams, consider tracking initiatives using these metrics & framework.



Why Is It So Hard to Tell If MLOps Is Working?



A Quick Tour: North Star Metrics





A Quick Tour: **OMTM**

OMTM Examples for Startups (< 50 employees)

- Number of website visitors per month
- Average click-through percentage from homepage to checkout
- Retention of new customers after 7 days

OMTM Examples for Scaleups (50-200 employees)

- Number of website visitors per month from SEO on our blog
- Conversion percentage of website visitors from Germany
- Retention of new, freemium customers after 7 days

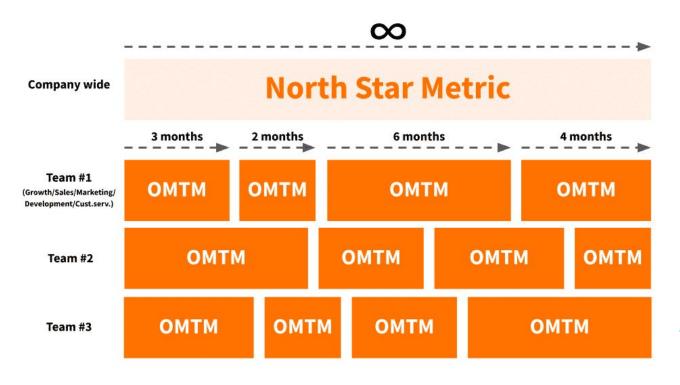
OMTM Examples for Corporates (> 200 employees)

- Number of website visitors / month via SEO around one specific topic
- Average CTR on the AddToCart Enterprise-Plan button on the Prices page
- Retention of new customers after experiencing the WOW moment

Source: "What is the One Metric That Matters?"



A Quick Tour: North Star + OMTM



Source: "What is the One Metric That Matters?"



A Quick Tour: KPIs & OKRs

KPIs vs OKRs: What's the Difference?

		Key Performance Indicators (KPIs)	Objectives and Key Results (OKRs)
WHAT	?	Numbers that track the operation of your business	Action-orientated goals and measures
FOUNDATION		Based on past results or future goals	Mission-based, aspirational and directional
DIRECTION	Щ	Monitors the "steady-state" and benchmarks	Audacious and bold, tied to mission
TRIGGERS	←• •→	Actions are prompted when numbers are off track	Actions are taken as issues arise
DURATION		Measured on an ongoing basis	Time-bound, often quarterly
LIFESPAN	Ō	May be the same from quarter to quarter, year to year	Change from quarter to quarter, year to year
What Matters			W

Source: "The Difference Between KPIs and OKRs"



Let's Get More Technical



A Quick Tour: **DORA Metrics...**



Source: "Humanitec -- Key
DevOps metrics to improve your
engineering setup"



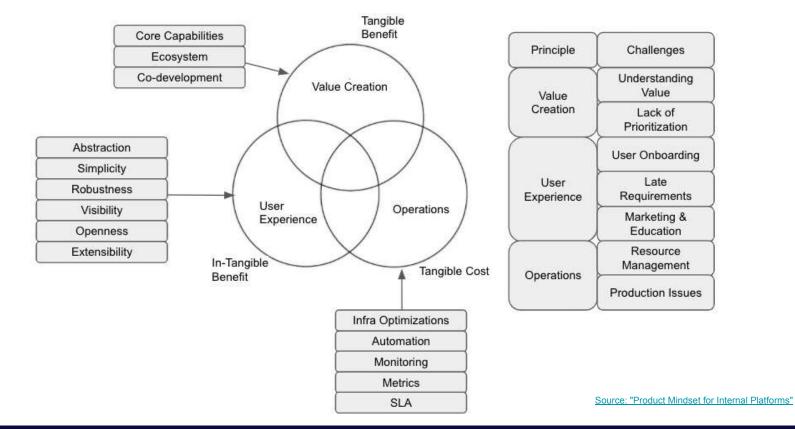
A Quick Tour: ... Risks, Throughput, Velocity



Source: "Humanitec -- Key DevOps metrics to improve your engineering setup"



A Quick Tour: SRE SLAs

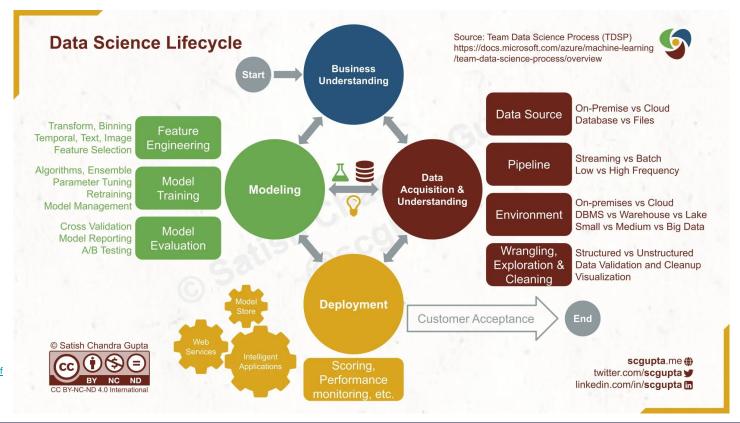




So It Should Be Easy, Right?



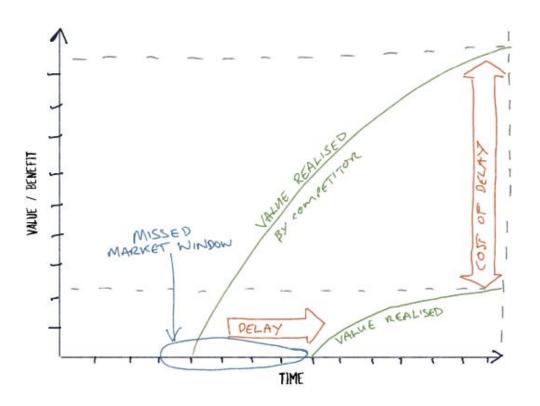
Challenge 1: **Defining ROI?**



Satish Gupta: "Survey of Machine Learning Lifecycle"



Challenge 1: Defining ROI?



Tom Geraghty: "Platform as a Product"



Challenge 1: **Defining ROI?**

It's the difference between:

- Outcome(If I did nothing)
- Outcome(I did something)

by a certain point-in-time



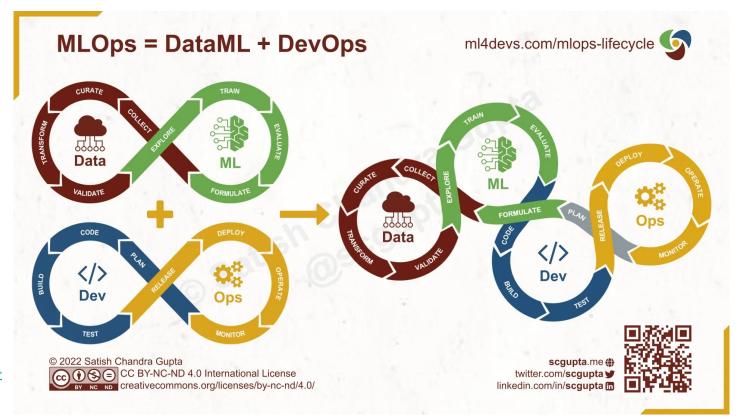
We Need To Fix Our **ROI Equation For MLOps Initiatives**



How Does The Data Science Experience Impact The BL?

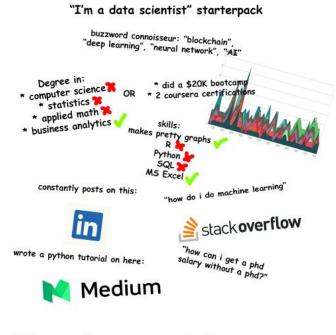


Challenge 2: Unclear Relationship

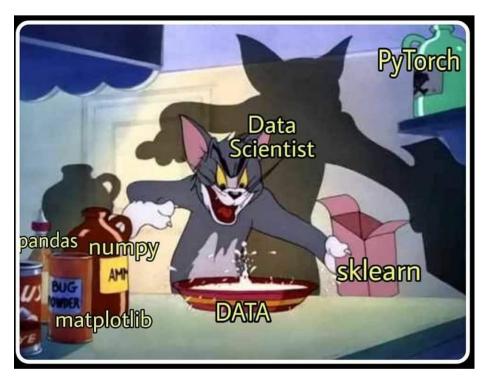


Satish Gupta:
"MLOps: Machine
Learning Life Cycle"

Challenge 2: Unclear Relationship



"I'm a data scientist" starterpack





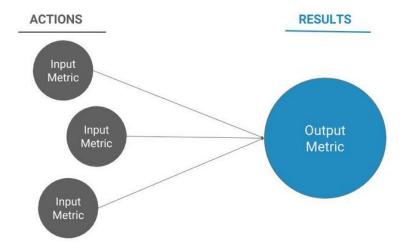
But Really, It's Our **Fault For Trying To Use Metrics OOTB**



Mistake 1: Vanity Metrics

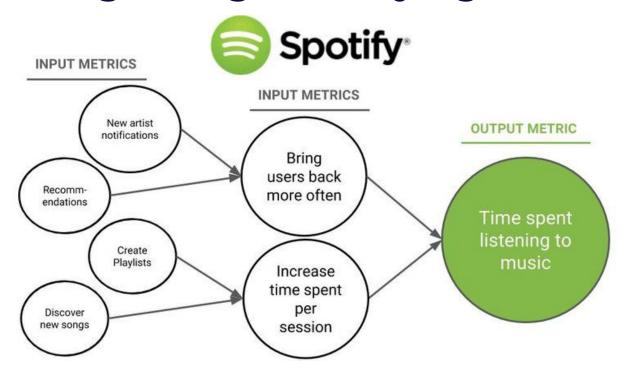
North star metrics (or OMTM) are output (results) based metrics, not input (action) based

Bain Balfour: "Don't Let Your North Star Metric Deceive You"





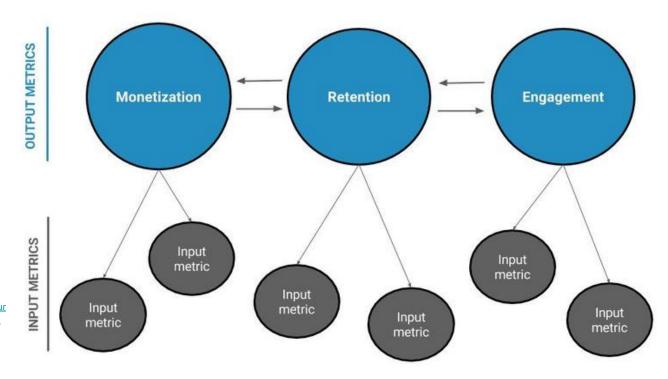
Mistake 2: Ignoring Underlying Behavior



Bain Balfour: "Don't Let Your North Star Metric Deceive You"



Mistake 3: Useless Simplification



Bain Balfour: "Don't Let Your North Star Metric Deceive You"



Mistake 4: Measuring Out Of Scope

Stop Trying To Measure Data Science Tasks As Part of MLOps

Measure Using Metrics Corresponding To The Right Layer



Mistake 5: Setting Targets Too Early

- DORA metrics are appropriate for checking vitals But they won't:
 - Tell you what you should be doing i.e. are they leading or are they lagging
 - Aren't appropriate for measuring implementation of a greenfield project



Mistake 5: **Setting Targets Too Early**

- Not specific enough to ML platforms & also not generalizable to be used as platform metrics across the board
- Also not useful as Developer Platform metrics
- Potentially also not good for a new ML project because of the experimental phase

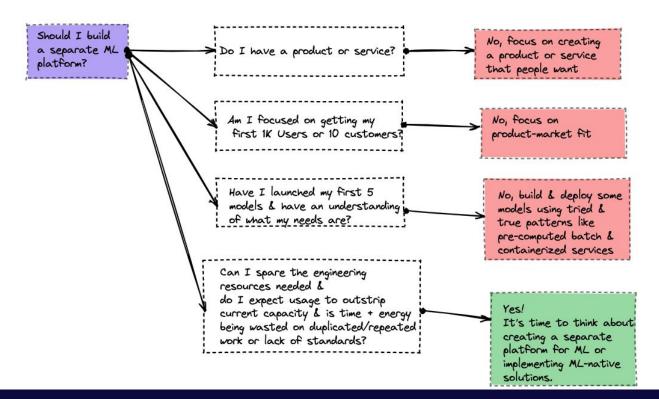


Here's The Game Plan



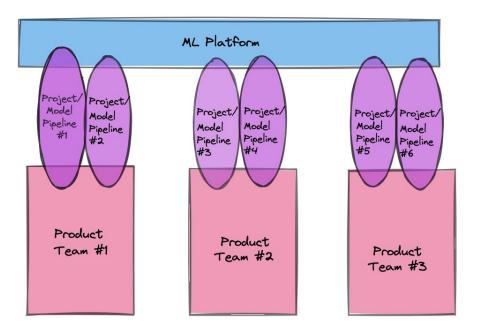


Step 1: Know Thyself





Step 2: Identify The Sandwich





Step 3: Decompose Sandwich Into Meaningful Layers

Platform Metrics



product metrics

(Accelerate metrics for platform services)

user satisfaction metrics

(Accelerate metrics for business services, NPS, etc)

adoption & engagement metrics

(% teams onboard, per platform and per service)

reliability metrics

(SLOs, latency, #Incidents, etc)

85





Step 4: Understand The Landscape

SPACE FRAMEWORK

S: Satisfaction and well being

P: Performance

A: Activity

C: Communication and collaboration

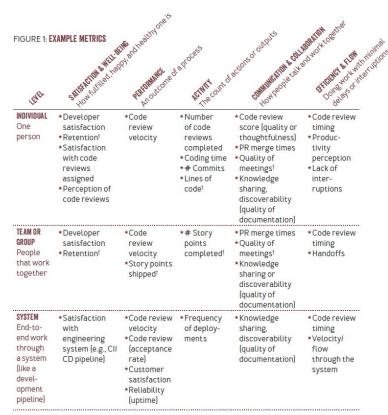
E: Efficiency and flow



Step 4:

Understand The Landscape

Paper: "The SPACE of Developer Productivity"



[†] Use these metrics with (even more) caution — they can proxy more things.



And Above All Else...



... Be Honest About Your **Goals & Desired** Outcomes ...



... Be Honest **About Your Current** Constraints...



... Ruthlessly Prioritize ...



... And Don't Be Afraid To Adapt To Change!



Goals

- What are the North Star metrics of an effective ML and MLOps practice?
- How should MLOps and Data Science teams be organized to provide the most value to the organization?
- How do we connect the tactical, operational,
 & strategic layers of an MLOps practice?



My BLUF

Don't Build An ML Platform Unless...

If You Do Build A Platform....

 When You Have a Platform & a Team, measure using these metrics...



Repository



bit.ly/3Yz2G95



Docs

bit.ly/423SE2W



LinkedIn



Mikiko Bazeley Head of MLOps

