adidas

SCALING DATA ANALYTICS DELIVERY MODEL

DEVOPS BY THE BOOK IN BI

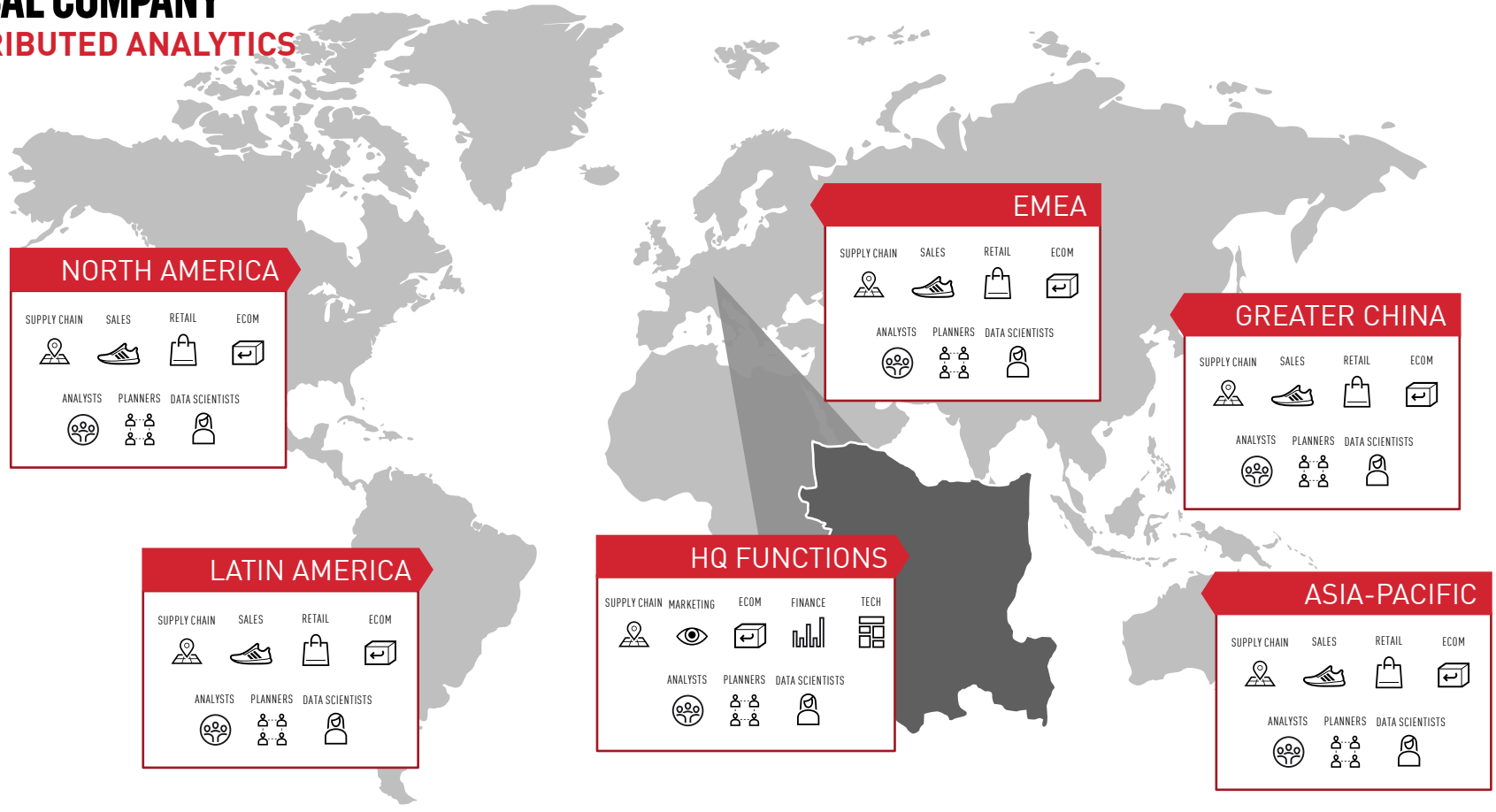
DMITRY LUCHNIK
DIRECTOR SOLUTION ARCHITECTURE – DATA & ANALYTICS

WHAT IS DATA ANALYTICS AT ADIDAS

GLOBAL COMPANY

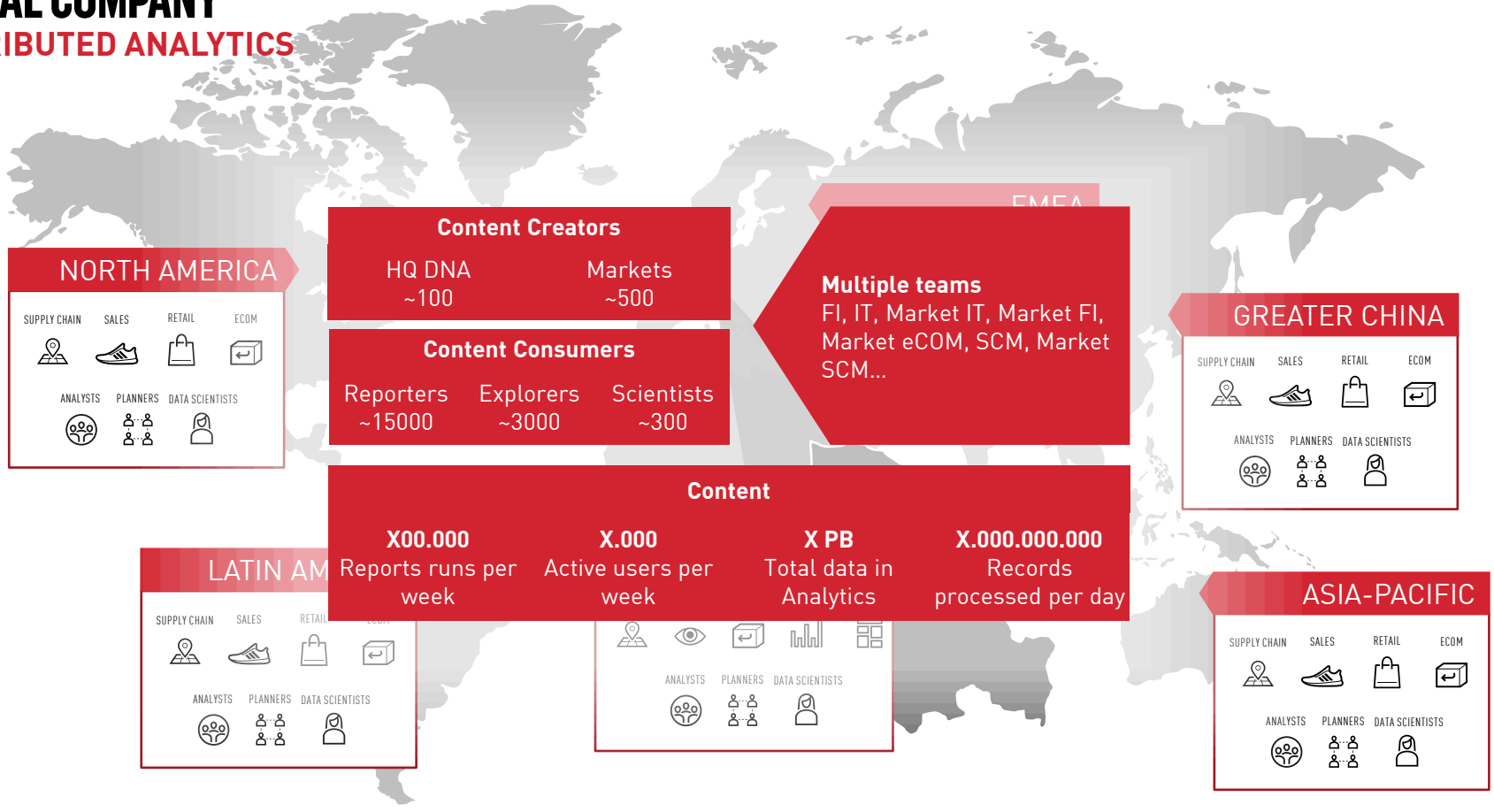
GLOBAL COMPANY

DISTRIBUTED ANALYTICS



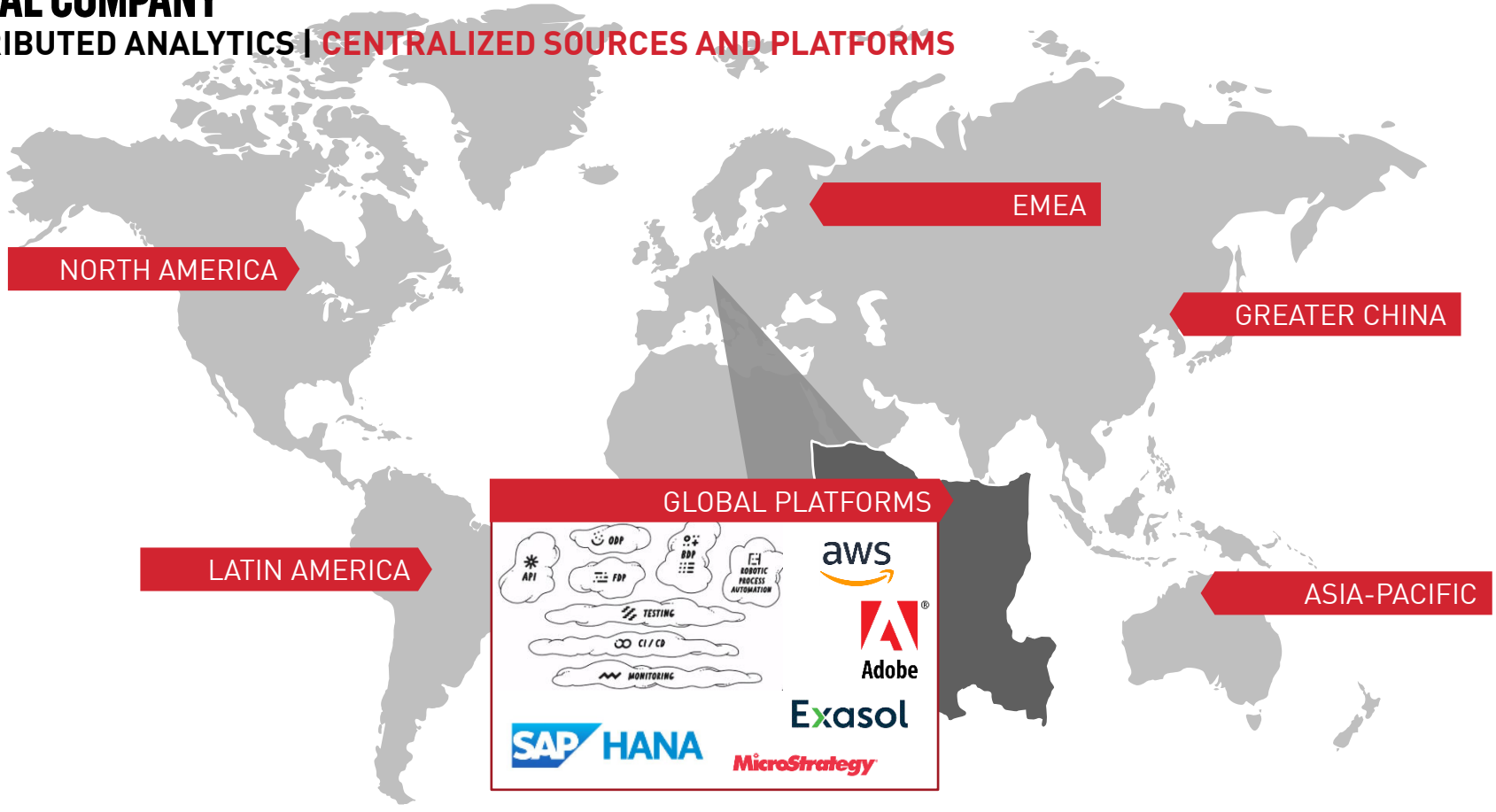
GLOBAL COMPANY

DISTRIBUTED ANALYTICS



GLOBAL COMPANY

DISTRIBUTED ANALYTICS | CENTRALIZED SOURCES AND PLATFORMS



DATA ANALYTICS OVERVIEW

DRIVING THE DIGITAL TRANSFORMATION TOWARDS DATA-DRIVEN DECISION MAKING

Data consumption tailored for different personas

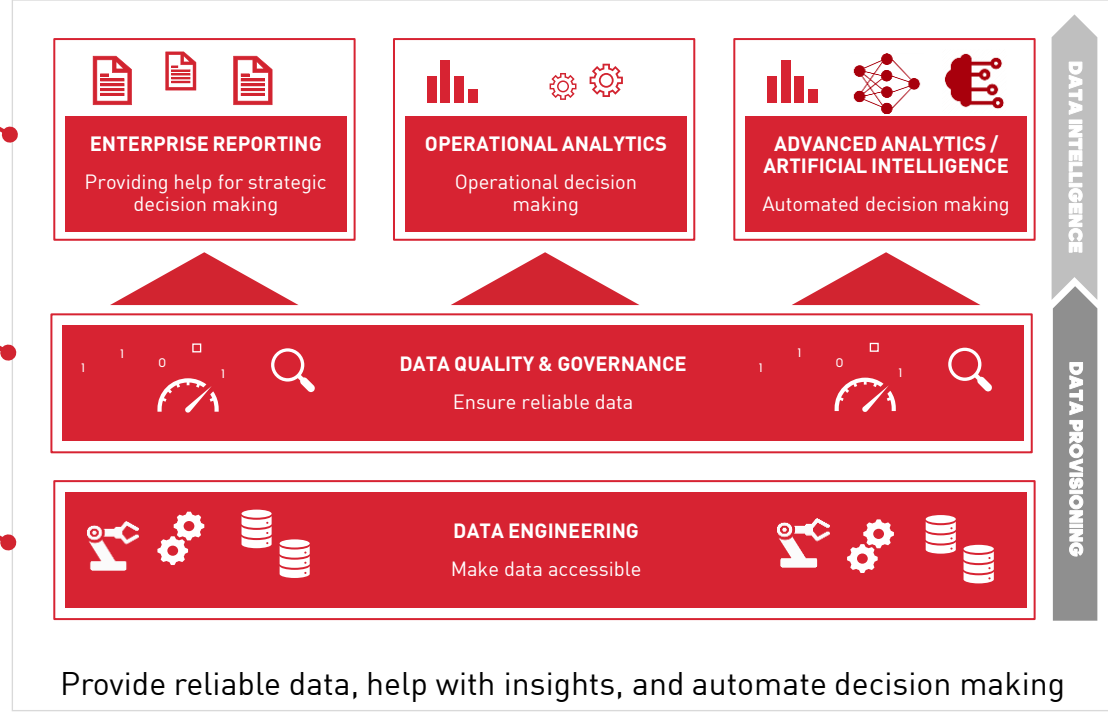
Multiple possibilities to interact with the data: standardized reports and dashboards, flexible data exploration, data labs.

Reliable data

Ensuring right and reliable data, data catalog, providing transparency on the data quality.

Connected data experience

Creating a connected data experience by making data accessible in an integrated enterprise-wide data lake. Providing tools and platforms for data exploration.



Provide reliable data, help with insights, and automate decision making

ZOOMING INTO OPERATIONAL ANALYTICS

WHAT IS OPERATIONAL ANALYTICS?

Enterprise Reporting
Standardized Reports and Dashboards
for
Enterprise Wide Performance Management

EXAMPLE

Operational Analytics
Data and Platforms
for
Operatioal Decision Making and Function / Market Reporting

EXAMPLE

Advanced Analytics
AI Algorithms
for
Semi automated Decision Making

EXAMPLE

TYPICAL OA QUESTIONS

- How to rebalance our stock between European countries?
- When it worth to do it with an air shipping?
- What are current deviations from standard lead times from factories to customers?
- How to balance load in factories?

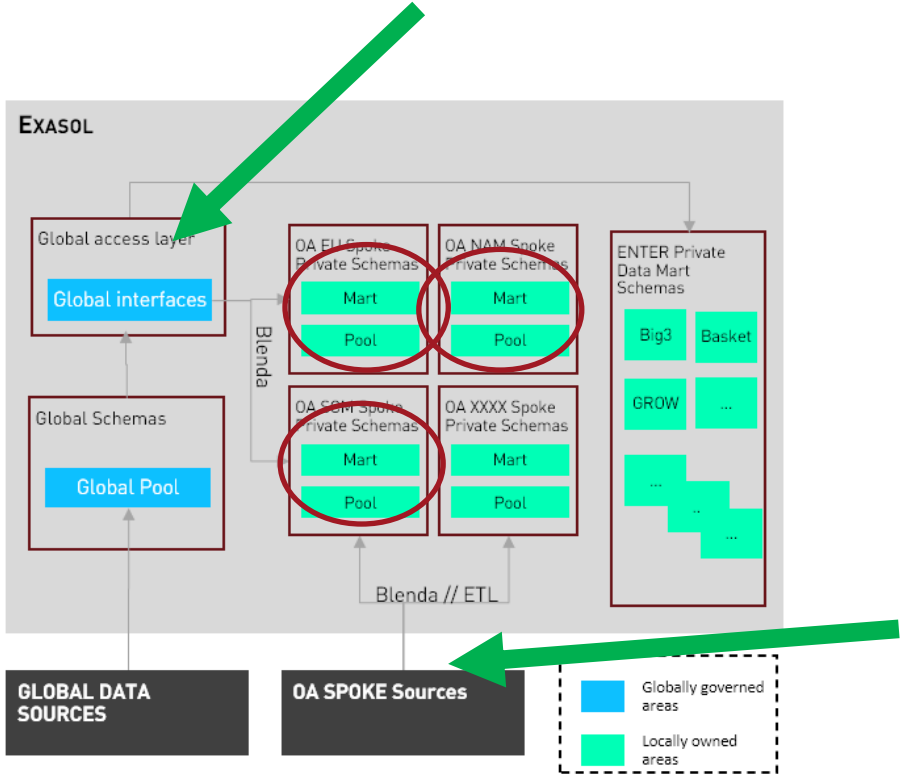
OA OPERATES IN “HUB AND SPOKES” MODEL

HUB – provides platforms, tools, ways of working and Global data

SPOKES – are empowered BI teams, located close to the analytics customers.

MANAGING OA DATA

Global data is **centrally** provisioned by the Global **Hub** team



To truly unlock potential of Local Analytics, it's often required to combine **Global** data with the **Local** addons.

Then transform and extract Local **meaning**.

Spokes are fully enabled to do it.

IS THERE A TECHNICAL CHALLENGE?

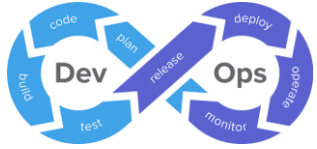
- Diverse infrastructure
- Multiple teams / many Markets
- Own processes
- Diverse skill levels

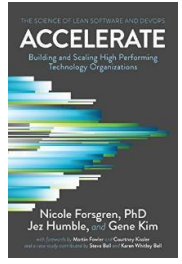
ARCHITECTURE FORMULA

TO ENABLE SPOKES AND ACCELERATE THEIR DELIVERY

DWH EXPERIENCE + **RESEARCH ON CI / DEVOPS CULTURE** + **ARCHITECTURE PATTERNS** → **BLENDA ARCHITECTURE**

Governance
Global/Local areas
Layers
How people work
Processes





Loosely coupling
Separation of Concerns
Transparency
Test Automation
Tests on every change
Deploy Automation
Automatic Reports
Notifications
Free Components:

- Bitbucket
- OpenSource
- Jenkins

Combining:

- Governance
- Speed
- Quality
- Empowerment

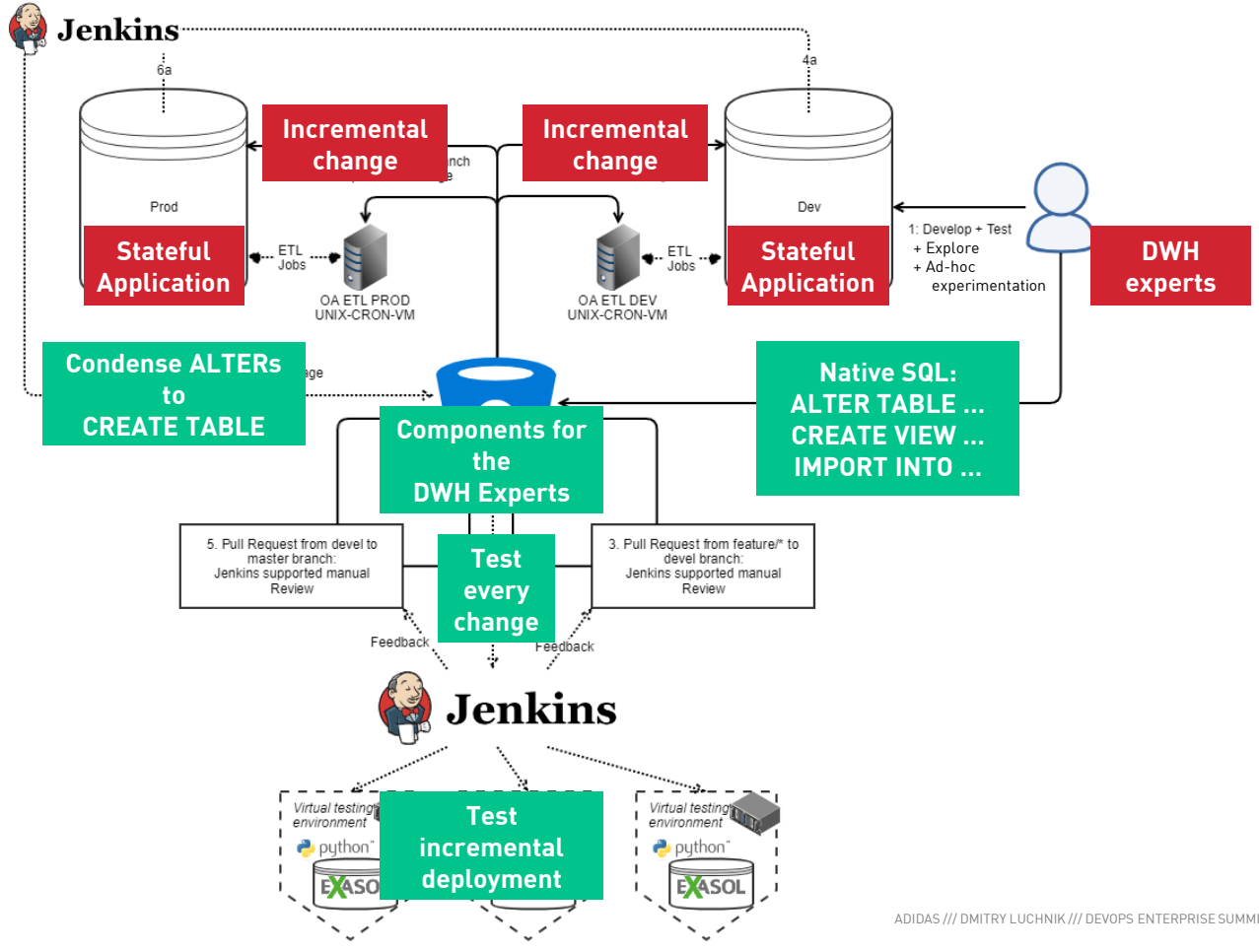
✖

Automation

↓

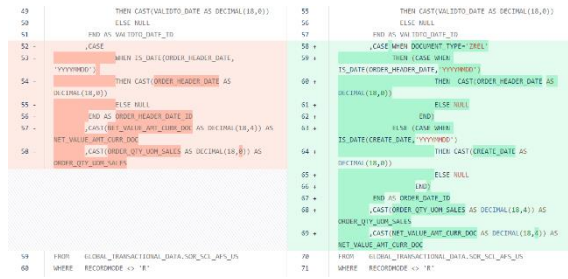
Acceleration

DEVELOPMENT AND DEPLOYMENT WITH BLENDA

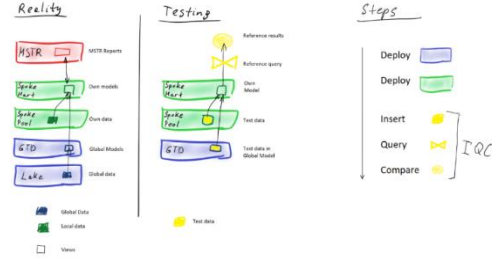


WHAT BLENDA TESTS?

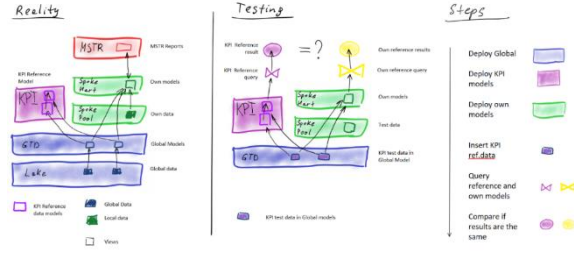
Emulate incremental production deployment



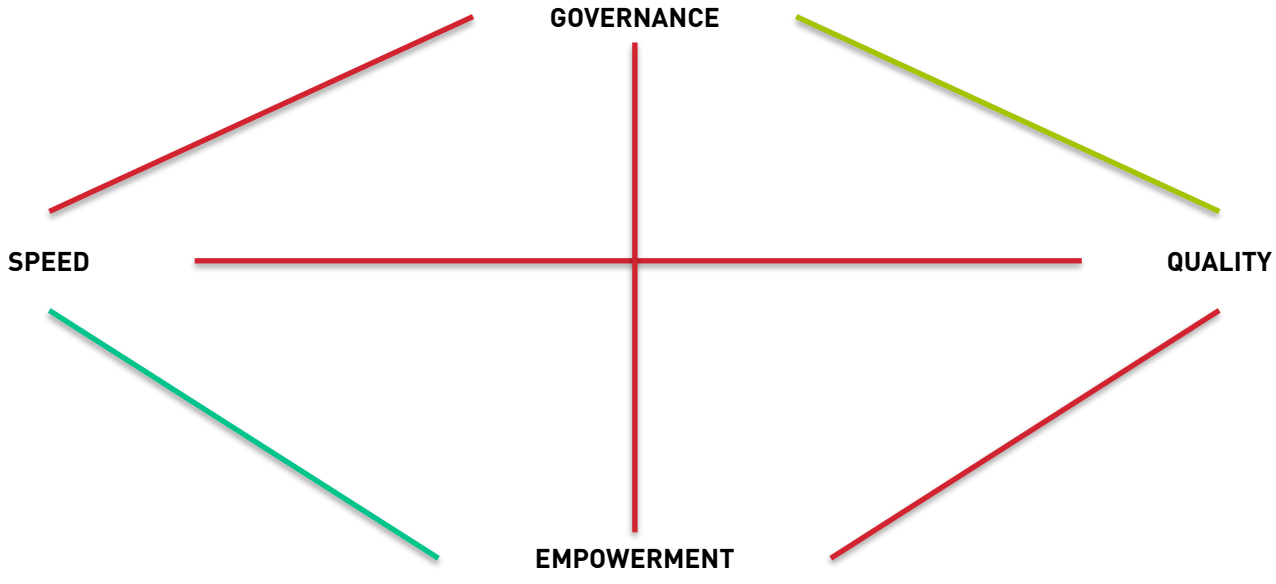
Impact to a frontend report



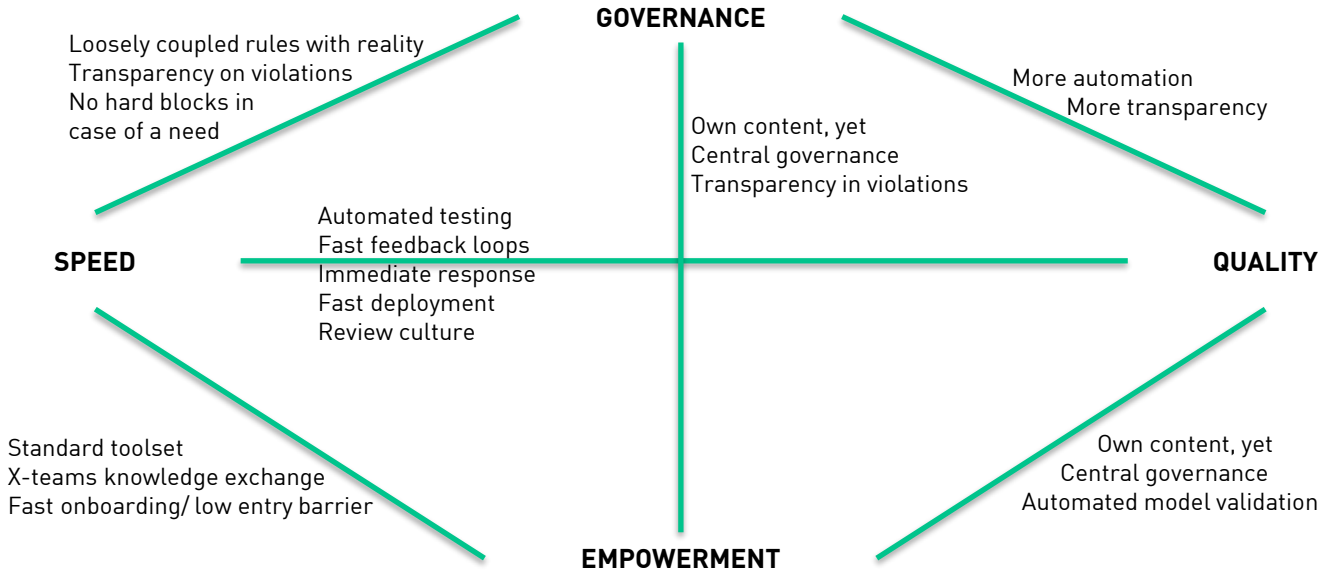
Alignment of a local KPI calculation with a global definition



WHAT USUALLY WORKS TOGETHER?

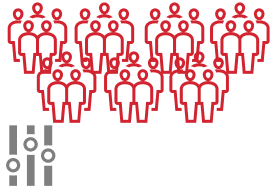
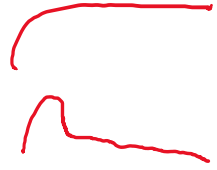
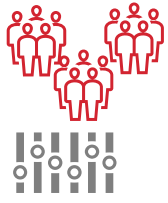
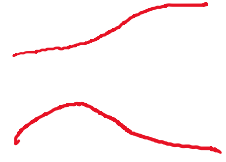


HOW BLENDA HELPS



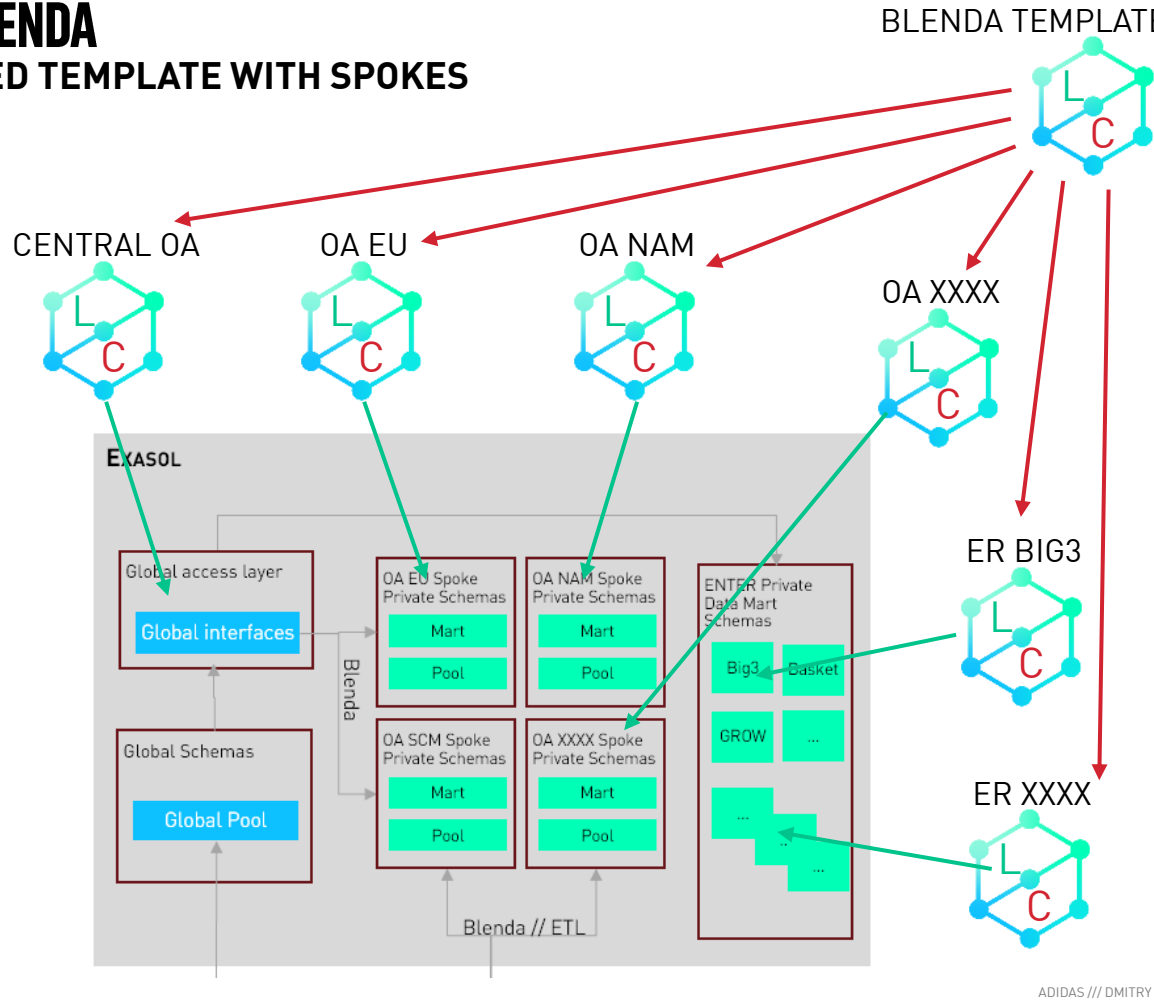
BLENDA INSIDE

CLEARLY SEPARATED LAYERS – LOOSELY COUPLED TO ENABLE SPEED

	LAYER	POPULATION & X-DOMAIN COMPLEXITY	UNDERSTANDING OF CONCEPTS AND TOOLS	LEARNING & INTRO CURVES	ONBOARDING
LOCAL OWNERSHIP	1 Spoke/ module content – data models, ETLs, etc.		DWH: ++++++++++ ETL: ++++++++++ SQL: +++++++++ CI: + git: +		From 1 hour
CENTRAL OWNERSHIP	2 QA framework – testing consistency, etc.		DWH: +++++++++++++ Testing: ++++ CI: +++ git: +++ python +		4h+
	3 Architecture, templates, deployment mechanics, technical tools		Architecture: +++++++++++++ CI/DevOps: ++++++++++ git: ++++++++++ python: ++++++++++ docker: +++++++		Constant improvements

CONSUMING BLENDA

LOOSELY COUPLED TEMPLATE WITH SPOKES



CENTRAL CONTENT

3 Framework, deployment mechanics, technical tools

2 Global tests

LOCAL CONTENT

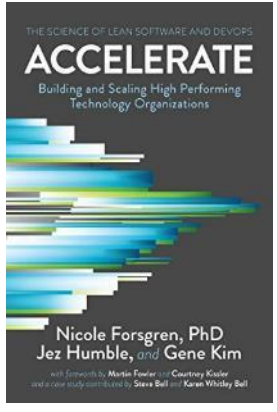
2 Local tests

1 Data models, ETLs

FOLLOWING THE BOOK...

CAPABILITIES FROM ACCELERATE

- ✓ Continuous Delivery
  - ✓ Version control
  - ✓ Deployment automation
  - ✓ Continuous Integration
  - ~~Trunk-based development~~
  - ✓ Test automation
  - ~~Test data management~~
  - ✓ Continuous Delivery
- ✓ Architecture
  - ✓ Loosely coupled
  - ✓ For empowered teams
- ✓ Process
  - ✓ Work in small batches.
  - ✓ Foster and enable team experimentation
- ✓ Lean management and monitoring capabilities
  - ✓ Have a lightweight change approval processes



BLENDA ENABLES

- Every update of a model, KPI or transformation is visible

  "Good, we've added just this one new calculation"

- Packaging of dependent changes in a single deployable block

  "OK! This ETL writes into a new column, and both changes are here."

- Testing of full data model on every change

  "My data model is always consistent."

- Automated on-demand deployments

  "CI looks good, pushing it now."

- Full internal know-how

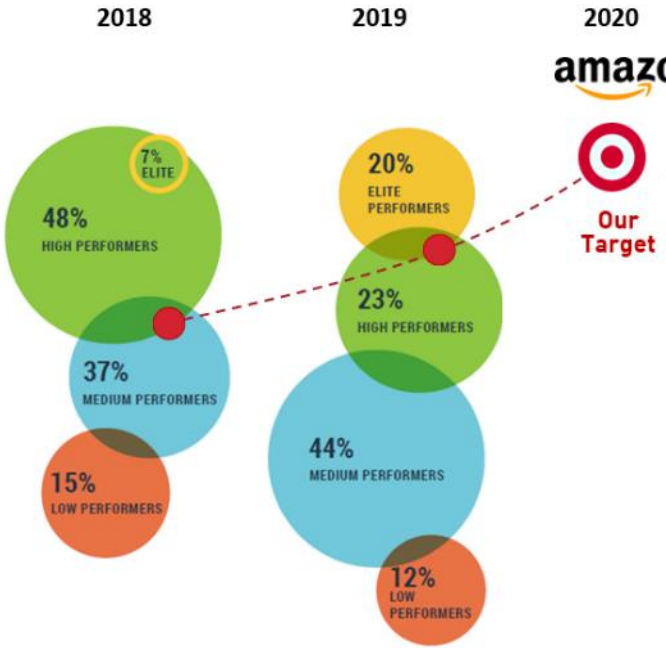
  "Ah! This is how that was done before!"

ADIDAS TECH'S TARGET IS TO BE IN THE GROUP OF ELITE PERFORMERS IN SOFTWARE DELIVERY

DATA WAREHOUSE & ANALYTICS CAN FOLLOW THE SAME PATH

INDUSTRY BENCHMARK – SOFTWARE DELIVERY PERFORMANCE

IN 2019, THE **% OF ELITE** PERFORMERS IN THE MARKET HAS ALMOST **TRIPLED**, DRIVING FURTHER **INCREASE OF CONSUMER EXPECTATIONS**



Source: DORA - State of DevOps 2019 sponsored by Google Cloud

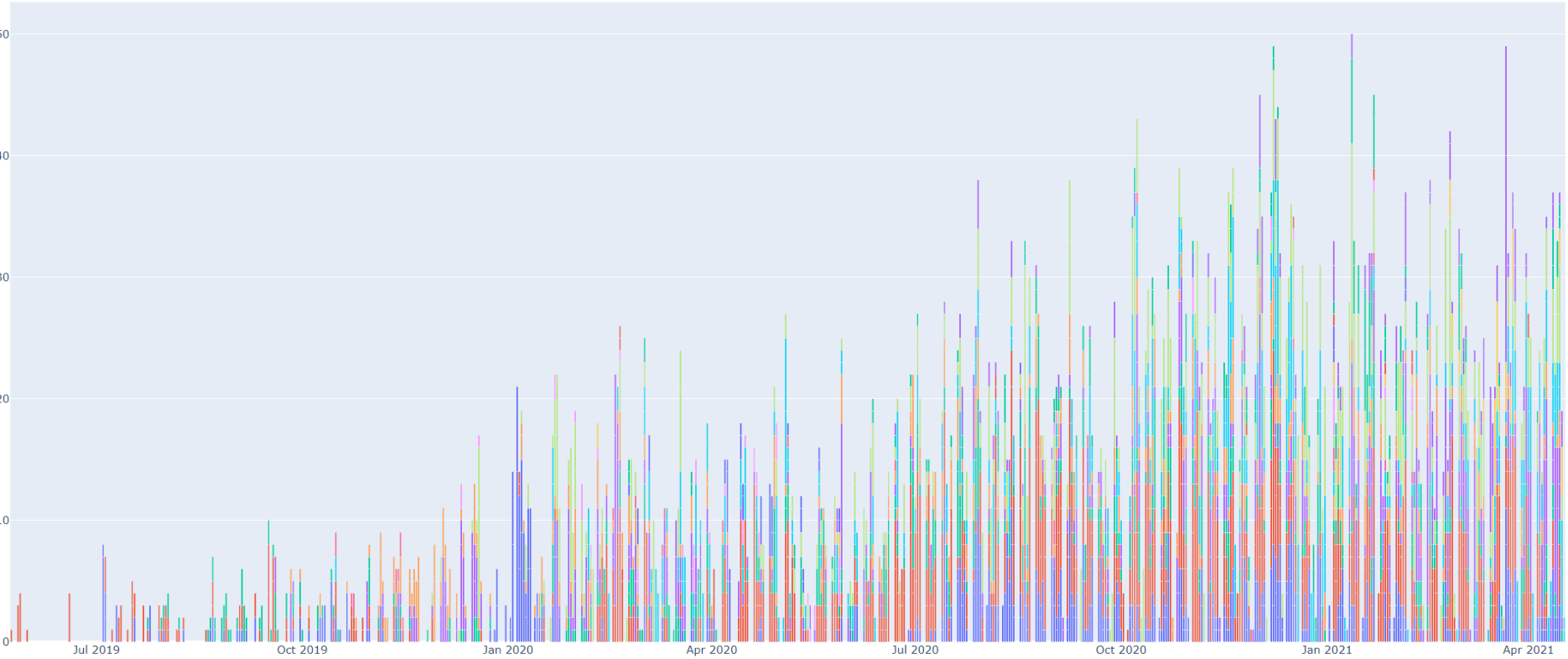
We are here

Aspect of Software Delivery Performance*	Elite	High	Medium	Low
Deployment frequency For the primary application or service you work on, how often does your organization deploy code to production or release it to end users?	On-demand (multiple deploys per day)	Between once per day and once per week	Between once per week and once per month	Between once per month and once every six months
Lead time for changes For the primary application or service you work on, what is your lead time for changes (i.e., how long does it take to go from code committed to code successfully running in production)?	Less than one day	Between one day and one week	Between one week and one month	Between one month and six months
Time to restore service For the primary application or service you work on, how long does it generally take to restore service when a service incident or a defect that impacts users occurs (e.g., unplanned outage or service impairment)?	Less than one hour	Less than one day*	Less than one day*	Between one week and one month
Change failure rate For the primary application or service you work on, what percentage of changes to production or released to users result in degraded service (e.g., lead to service impairment or service outage) and subsequently require remediation (e.g., require a hotfix, rollback, fix forward, patch)?	0-15%[b,c]	0-15%[b,d]	0-15%[c,d]	46-60%

1

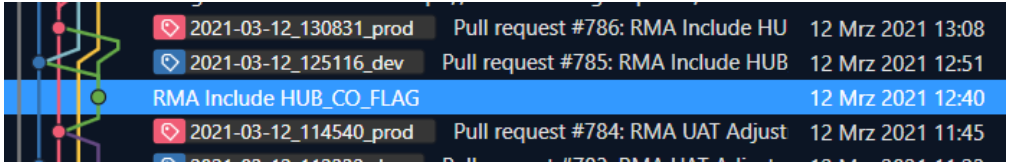
EFFICIENCY VIA BLENDA: DEPLOYMENT FREQUENCY

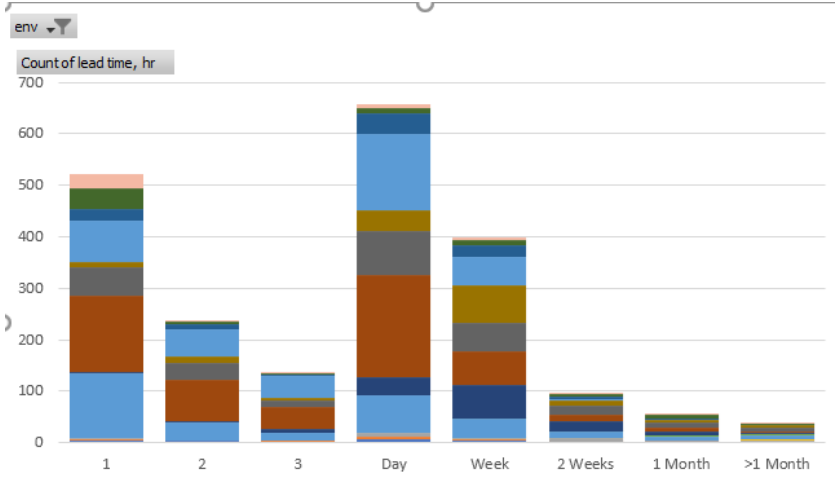
DEPLOYMENTS PER DAY

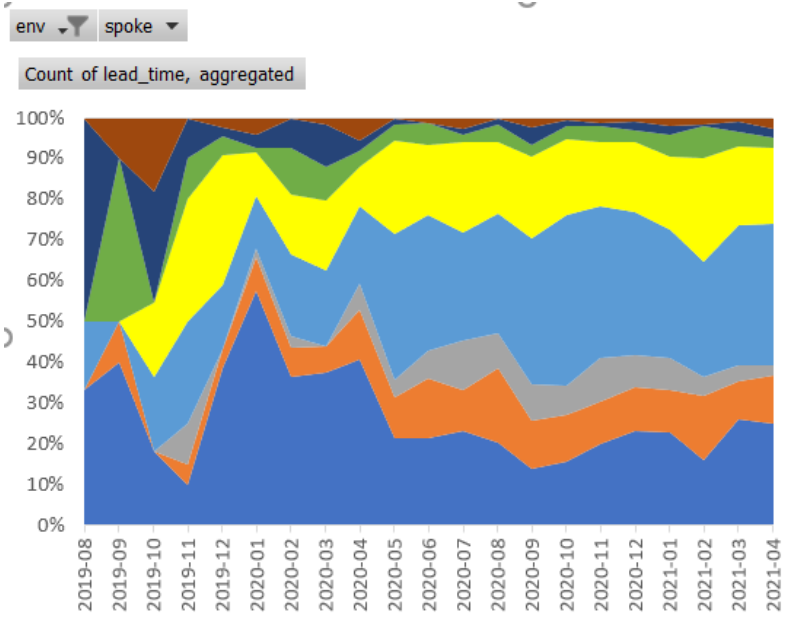


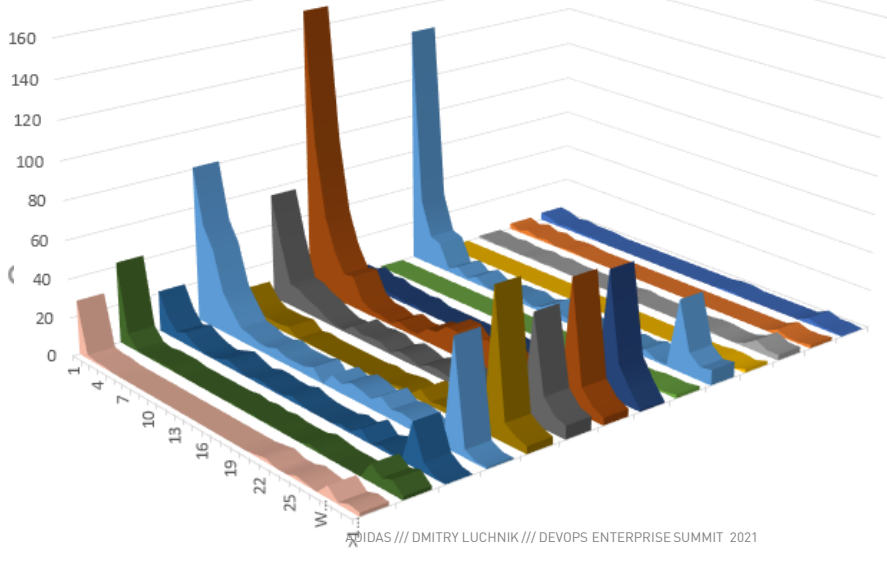
** every bar - is a number of deployments per day*
** every color – is an independent team*

EFFICIENCY VIA BLENDA: LEAD TIMES
SHORT *TO-PROD* CYCLES CREATE FLOW





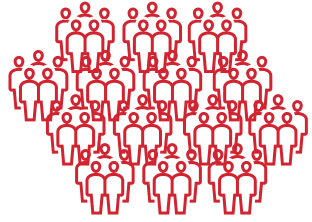




RESULTS IN A NUTSHELL

80Tb

100.000



3.000

20 mins

LEARNINGS

- **Native SQL** over abstractions like **liquibase** or **sql alchemy**
- Reliable **deployment mechanics** is a key
- **Fast CI** is a must

THE HELP I'M LOOKING FOR....

- How to speed up and scale processes for bringing new data to the Analytics ecosystem?

THANK YOU!

Dmitry Luchnik
Director Solution Architecture – Data & Analytics
adidas

dmitry.luchnik@adidas.com
www.linkedin.com/in/louchnik