Open for Innovation
KNIME

Data Science in Action: Stakeholder Collaboration

Paolo Tamagnini
Phil Winters

& the KNIME Team

Collaboration

Two or more people working together to complete a task or achieve a goal

Structured

Unstructured

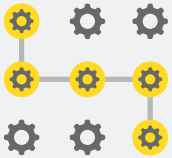
© 2020 KNIME AG. All rights reserved.

Stakeholders involved in Data Science…..

Team Leader
Data Science Practice

Data Engineers
Data Scientists
“Code First” / Package First

Business Analyst
Interactive User
Information Consumer



Model / ML Operations

IT Operations
Scalability and control
Centralized resources / strategies
Security

Compliance
Financial / risk oversight
Costs allocation
Governance, traceability, GDPR

© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

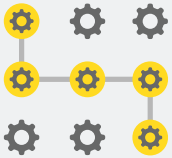
Successful Team Collaboration…..

Users

Team Leader
Data Science Practice

Data Engineers
Data Scientists
“Code First” / Package First

Business Analyst
Interactive User
Information Consumer



Model / ML Operations

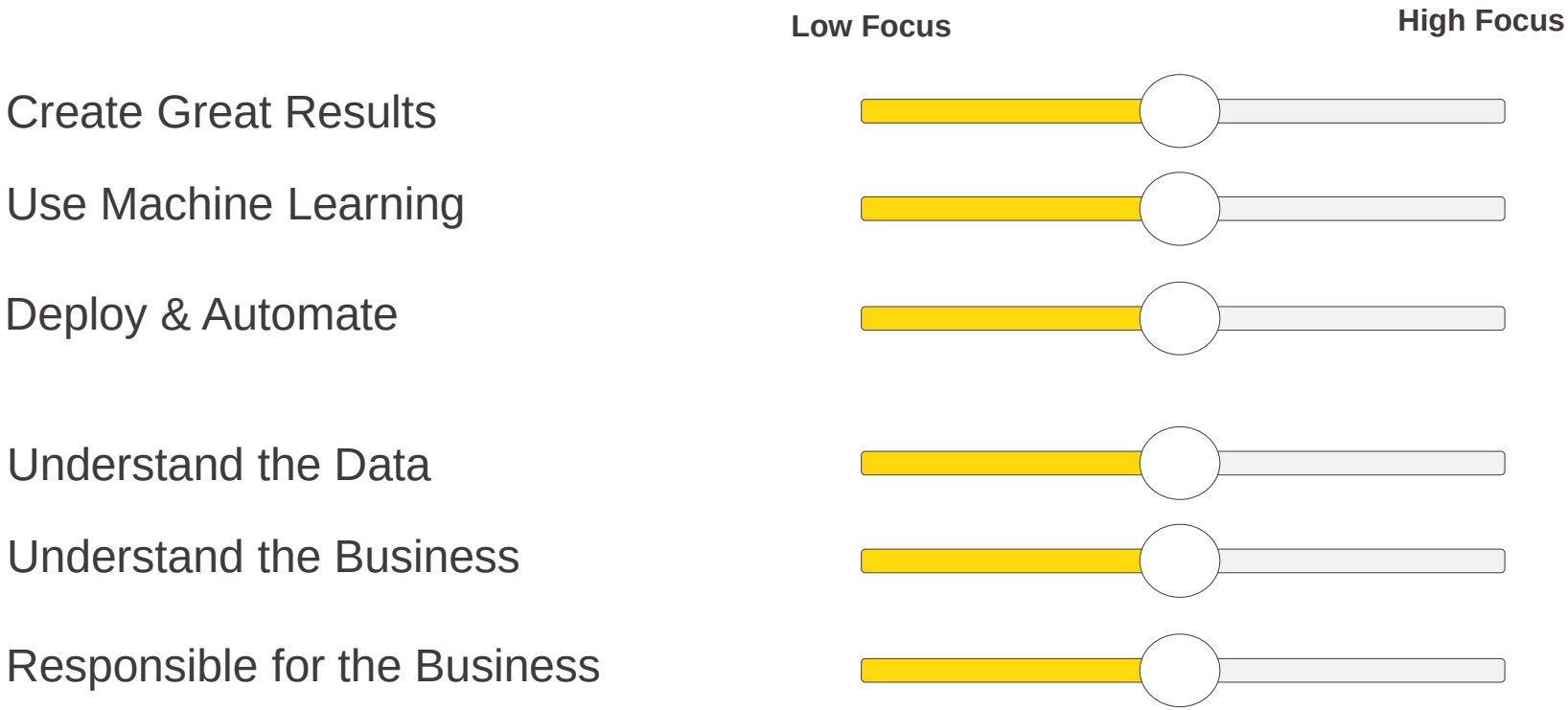
IT Operations
Scalability and control
Centralized resources / strategies
Security

Compliance
Financial / risk oversight
Costs allocation
Governance, traceability, GDPR

© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

Common across ALL Users, just different emphasis



© 2020 KNIME AG. All rights reserved.

Data Engineers / Data Scientists

Phil

Kathrin

SAME, only a difference in ML Focus!

Create Great Results

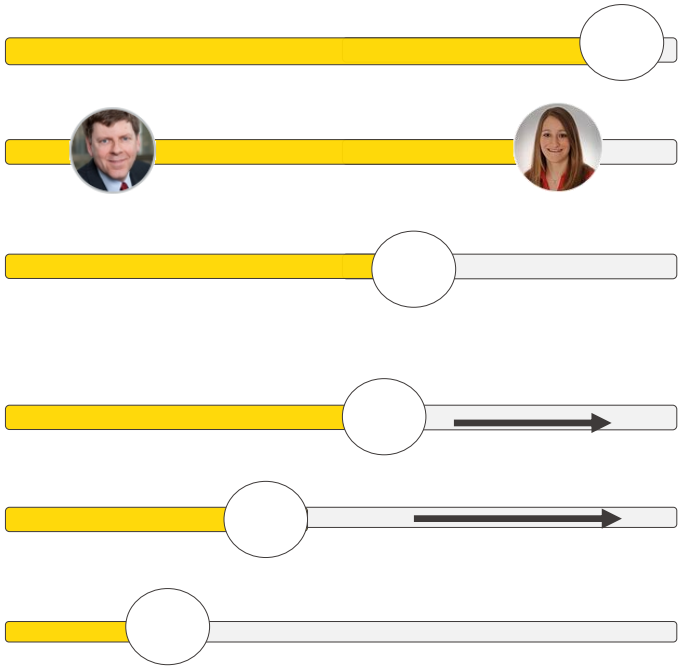
Use Machine Learning

Deploy & Automate

Understand the Data

Understand the Business

Responsible for the Business

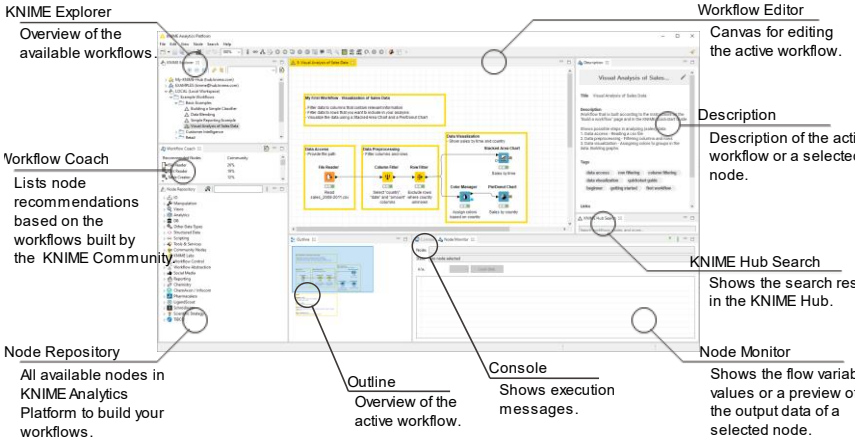


© 2020 KNIME AG. All rights reserved.

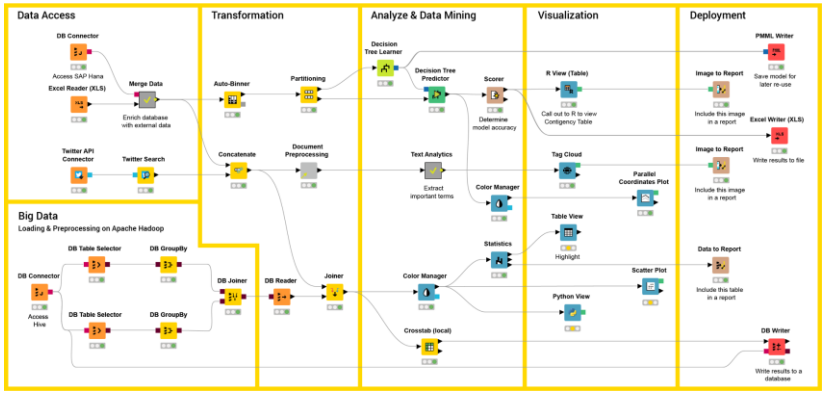
Open for Innovation KNIME

KNIME Provides

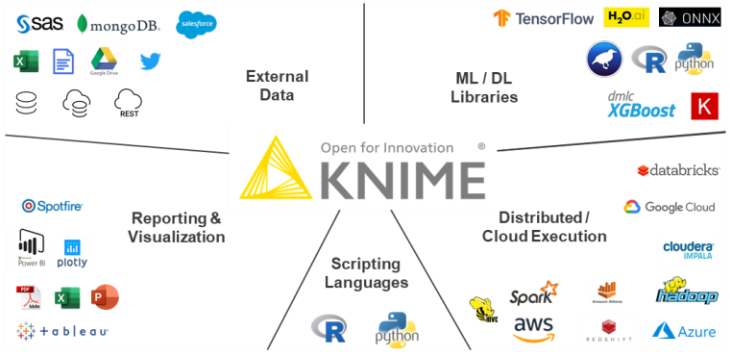
KNIME Workbench



3500+ Nodes



100+ extensions



© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

Data Engineers / Data Scientists / “Code-First”

Phil

Kathrin

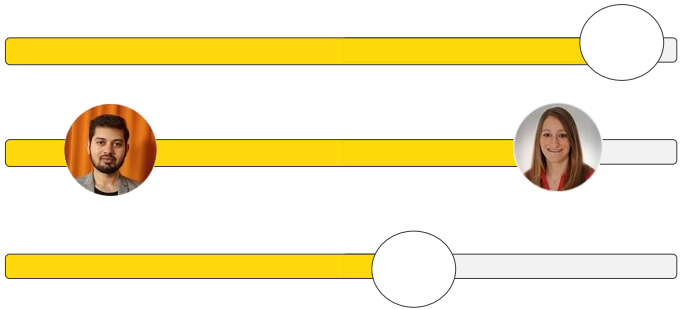
Monty

No Difference! Only preference
(Python in this case !)

Create Great Results

Use Machine Learning

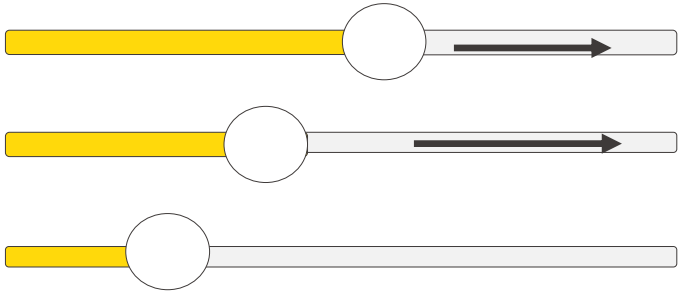
Deploy & Automate



Understand the Data

Understand the Business

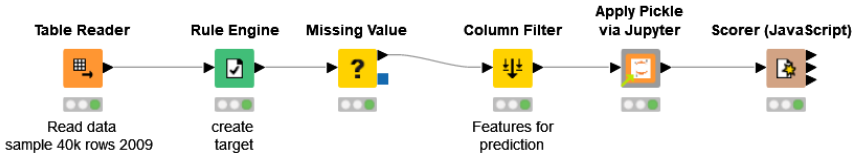
Responsible for the Business



© 2020 KNIME AG. All rights reserved.

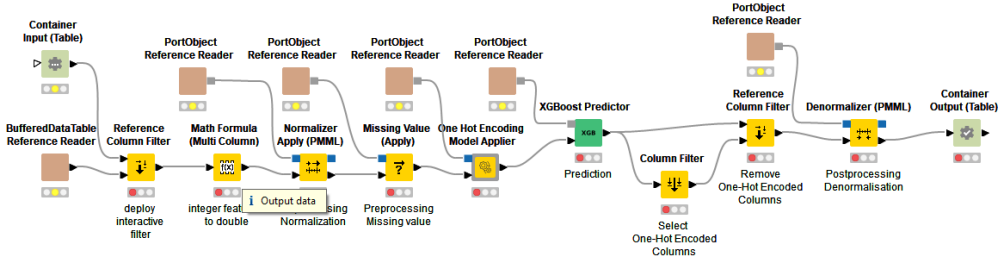
KNIME Provides: IDEs and Code-First Specifics

- Using Jupyter Notebooks in KNIME



- Using KNIME in Jupyter Notebooks

```
In [1]: import pandas as pd
        import knime
        knime.executable_path = "c:/Users/glandrum/Desktop/KNIME_3.7/knime.exe"
```



© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

Model / ML Operations

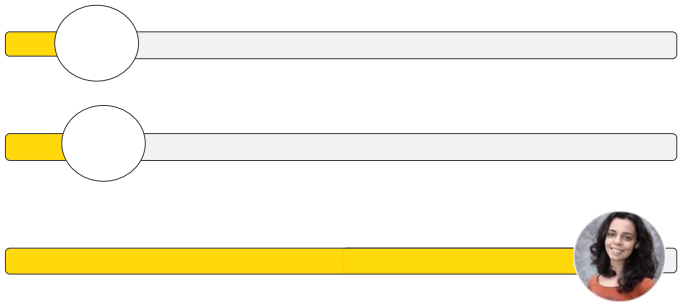
Ana

Deploy, otherwise it is useless…

Create Great Results

Use Machine Learning

Deploy & Automate



Understand the Data

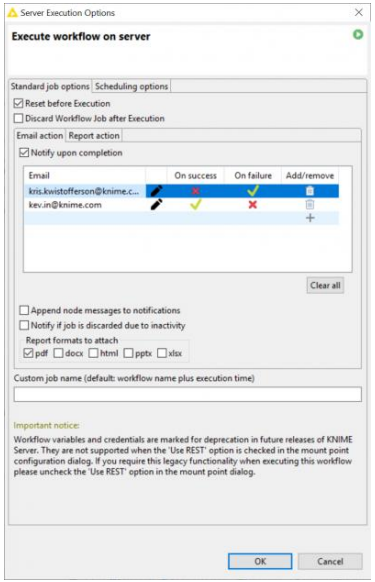
Understand the Business

Responsible for the Business

© 2020 KNIME AG. All rights reserved.

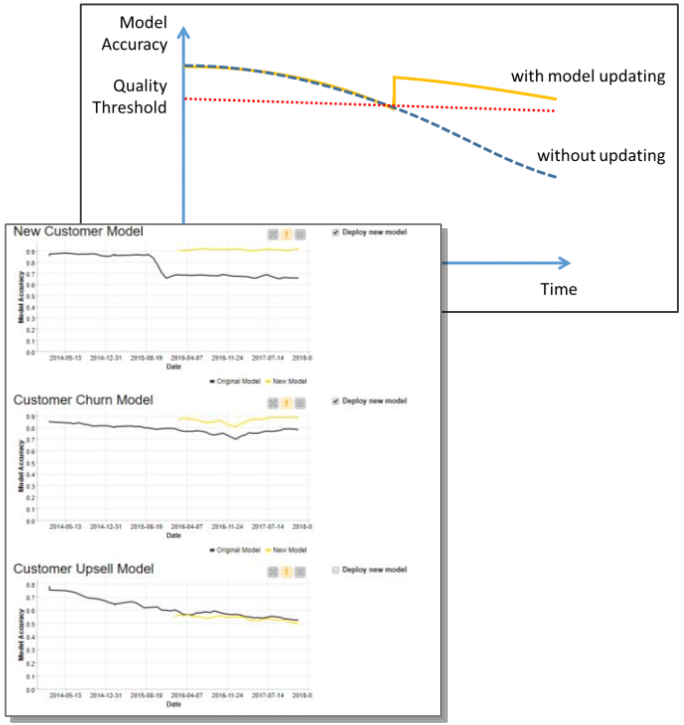
KNIME Provides: Automation Options

Automated workflow execution



- Scheduled
- Triggered/Manual Model Called
- (Rest / SAAS)
- Champion / Challenger
- Model Monitoring
- Retraining
- View, edit, execute workflows remotely

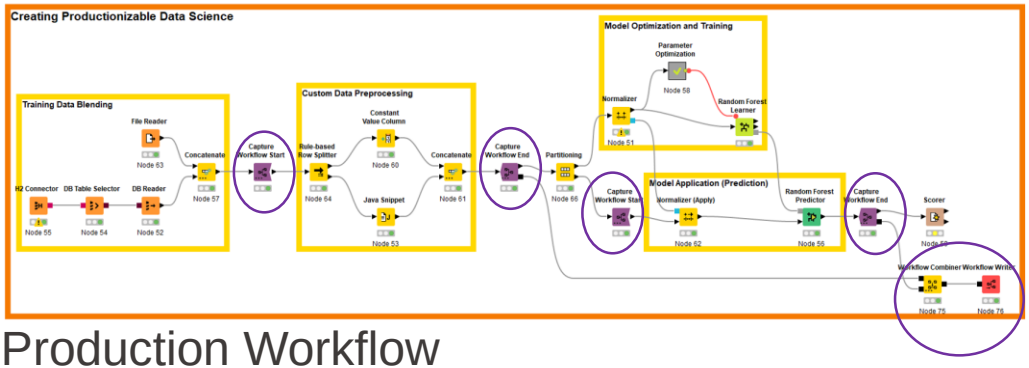
KNIME Model Process Factory



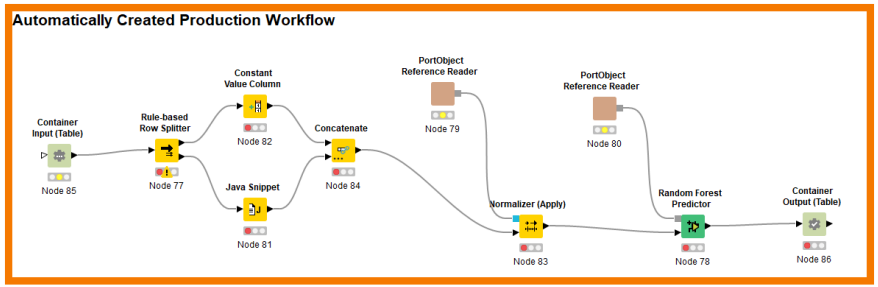
© 2020 KNIME AG. All rights reserved.

KNIME Provides: Integrated Deployment

Creation Workflow



Production Workflow



- Eliminates the gap between Creation and Production
  - mark what's necessary for production
  - not just the model
- no manual intervention
  - No rewrite, copy paste
- Always in sync

© 2020 KNIME AG. All rights reserved.

Successful Team Collaboration…..

Team Leader
Data Science Practice

Data Engineers
Data Scientists
“Code First” / Package First

Business Analyst
Interactive User
Information Consumer

Model / ML Operations

IT Operations
Scalability and control
Centralized resources / strategies
Security

Compliance
Financial / risk oversight
Costs allocation
Governance, traceability, GDPR

© 2020 KNIME AG. All rights reserved.

Information Consumer / Interaction User

Heather

Christiane

Use workflows……but never see them.

Create Great Results

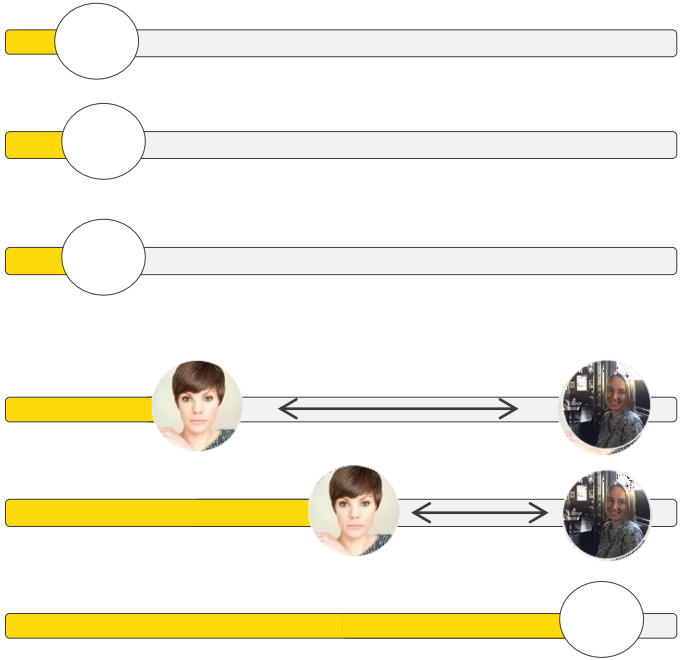
Use Machine Learning

Deploy & Automate

Understand the Data

Understand the Business

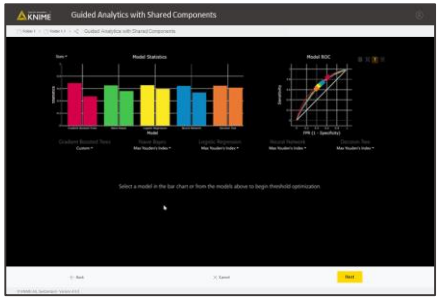
Responsible for the Business

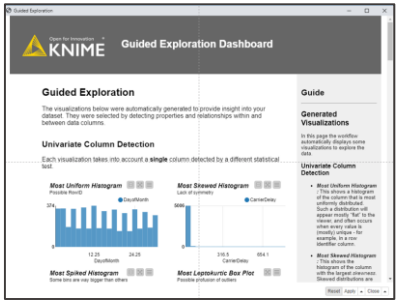


© 2020 KNIME AG. All rights reserved.

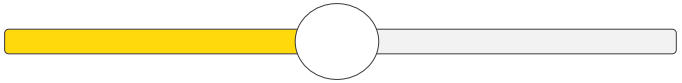
KNIME Provides: Guided Analytics

- Appropriate levels of Automation & Human Interaction depending on **task** and **audience**

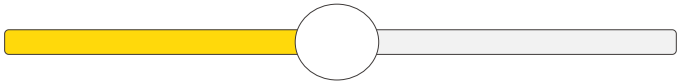




Automation



Interaction



© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

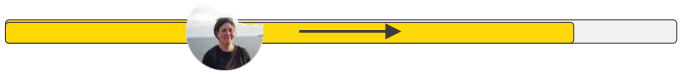
Business Analyst

Daria

Create Great Results

Use Machine Learning

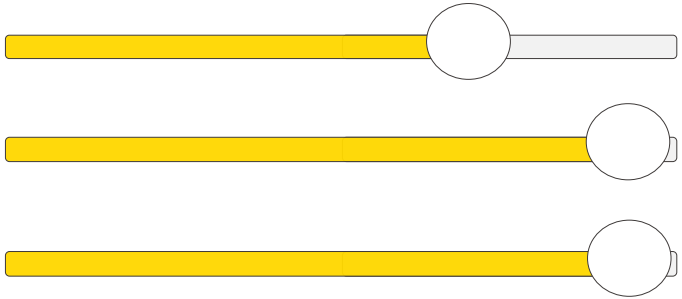
Deploy & Automate



Understand the Data

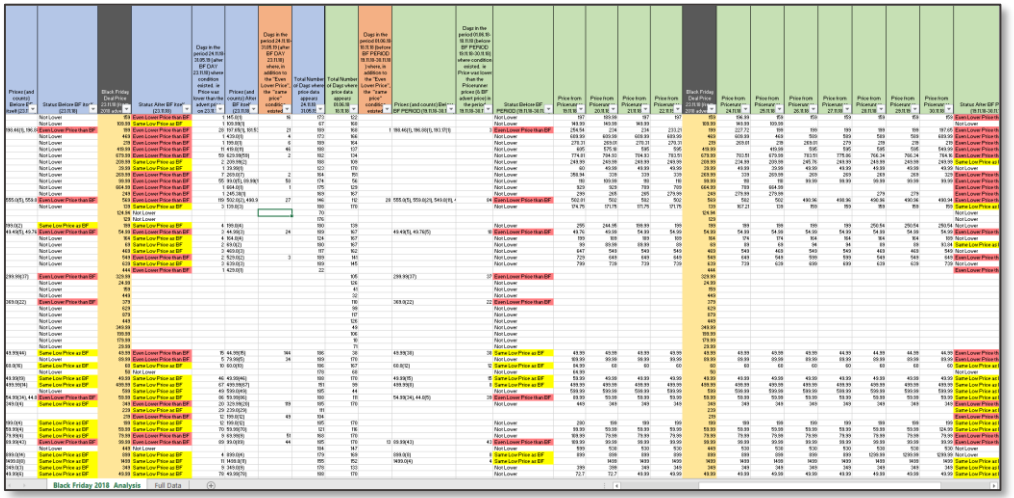
Understand the Business

Responsible for the Business



© 2020 KNIME AG. All rights reserved.

KNIME Provides: Integrate with front end of choice



- EXCEL Exploitation*
  - Not just CSV Files…
- Power BI
- Tableau
- Spotfire
- Qlik

*Thank You Continental !

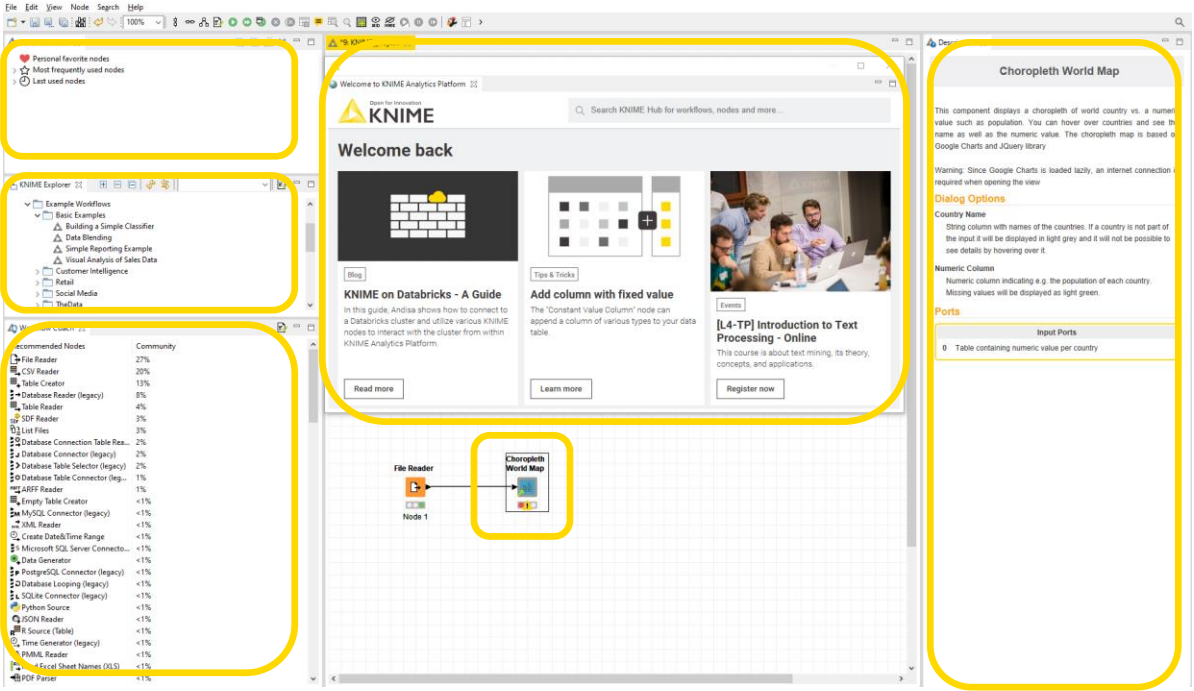
https://www.knime.com/community/continental-nodes-for-knime

© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

KNIME Provides: Tailored KNIME Analytics Platform

Surface capabilities so users can appropriately create workflows themselves



Tailor / Focus:

- Welcome
- Help /guidance
- Available Workflows
- Node availability
- Components & Descriptions
- Workflow coach for recomendations

© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

DS Practice Leader

Rosaria

Leverage, Exploit, Grow, Limit Risk, Satisfy

Create Great Results

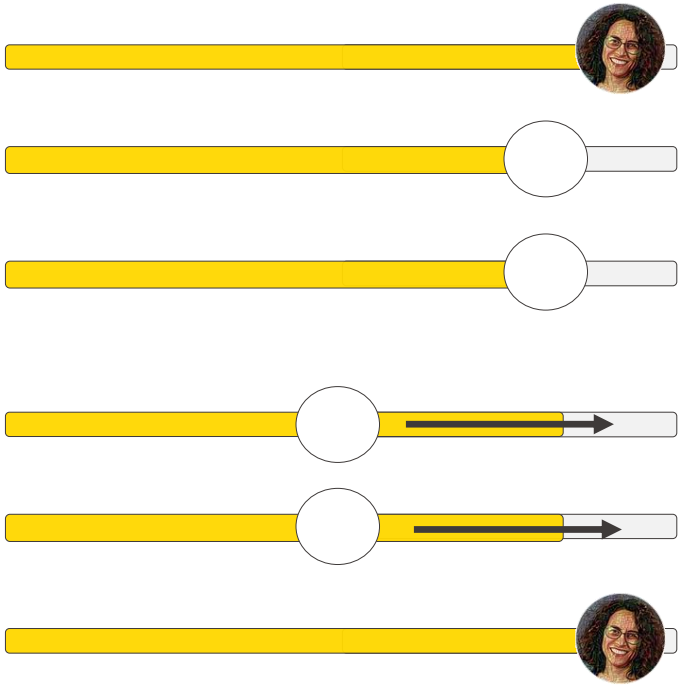
Use Machine Learning

Deploy & Automate

Understand the Data

Understand the Business

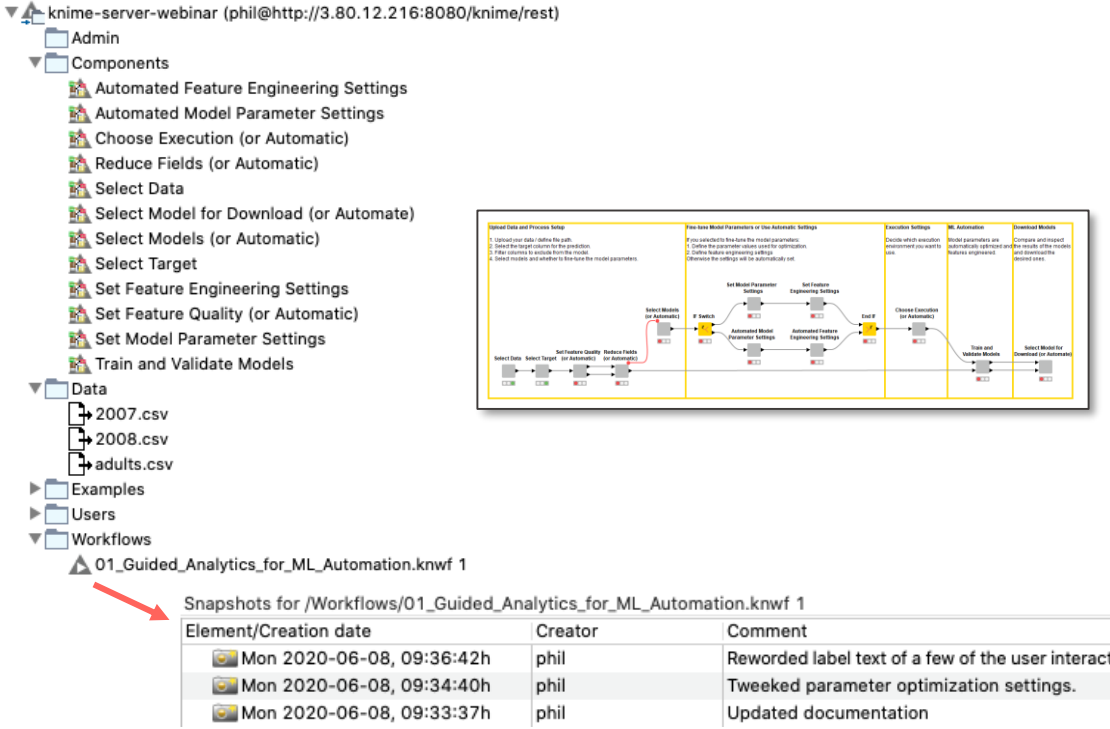
Responsible for the Business



© 2020 KNIME AG. All rights reserved.

Captured, Instantiated, Secure, Versioned, Reproducible

Workflows and Components



KNIME Provides:

- Instantiation/ Updating
- Security (create, use, etc.)
- Versioning and snapshots
  - Version comparison
- Locking / Encryption
- Reproducability
- Guaranteed backward compatible

© 2020 KNIME AG. All rights reserved.

Explainability / Interpretability

KNIME Provides:

- Methods
- Examples
- Interactive Components
  - XAI

© 2020 KNIME AG. All rights reserved.

Successful Team Collaboration Uses The KNIME Platform

Team Leader
Data Science Practice

Data Engineers
Data Scientists
“Code First” / Package First

Business Analyst
Interactive User
Information Consumer

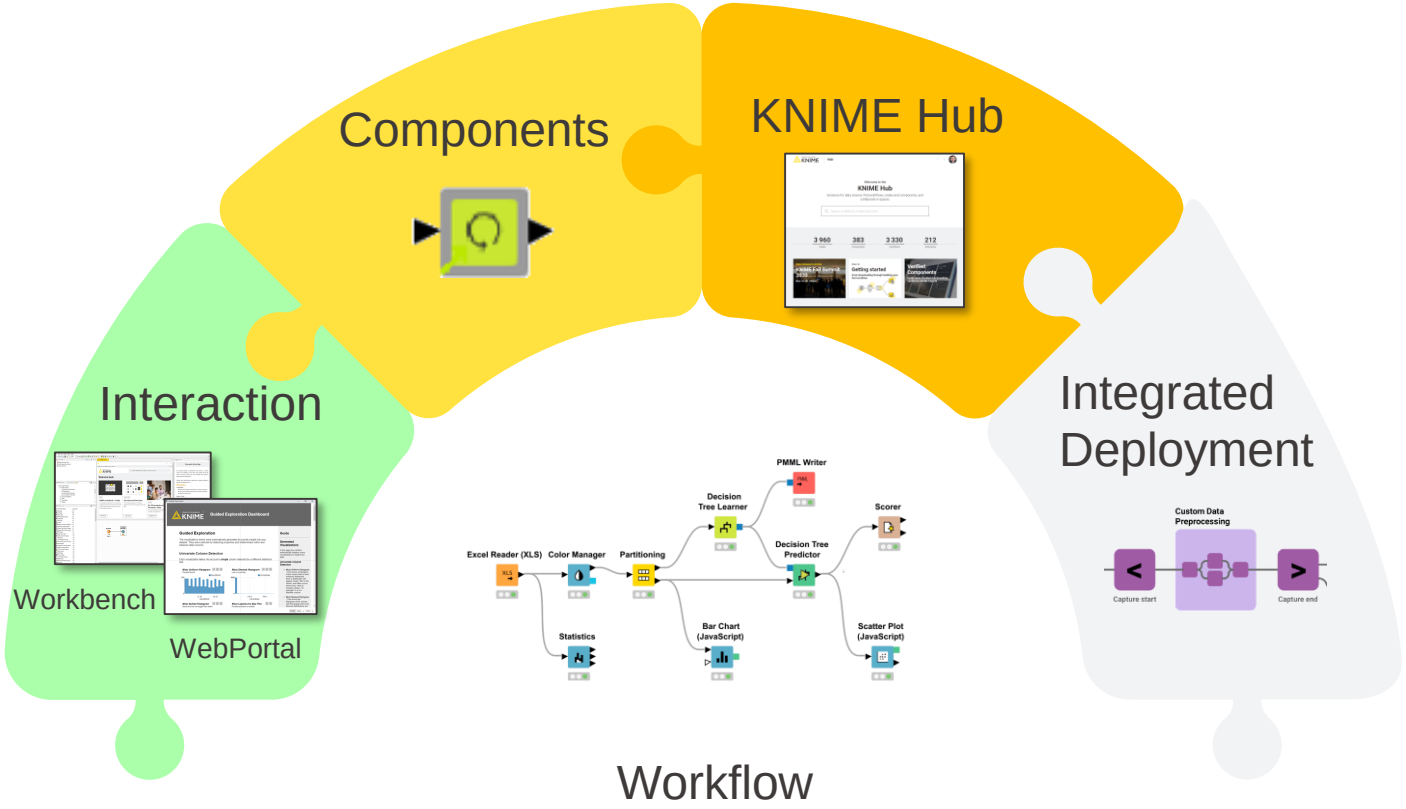
Model / ML Operations

IT Operations
Scalability and control
Centralized resources / strategies
Security

Compliance
Financial / risk oversight
Costs allocation
Governance, traceability, GDPR

© 2020 KNIME AG. All rights reserved.

Highlights for Team Collaboration



© 2020 KNIME AG. All rights reserved.

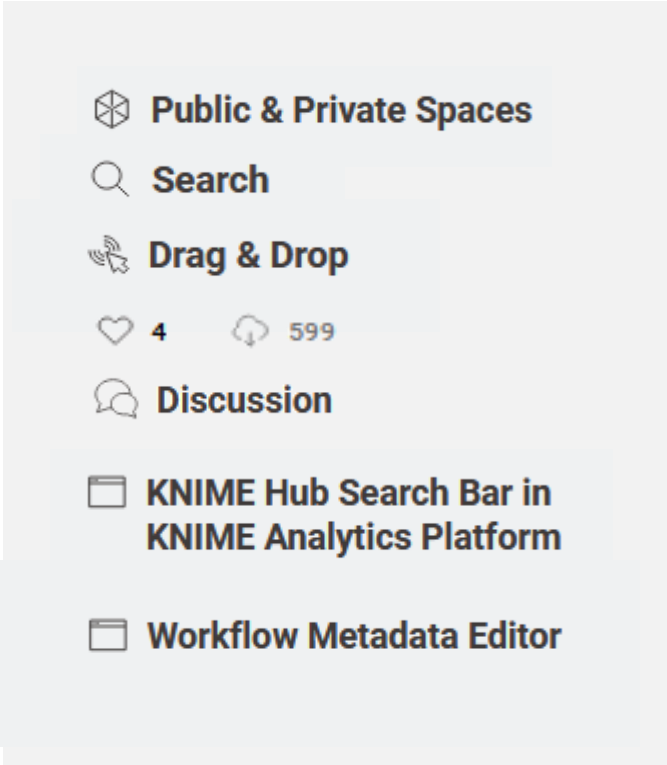
Open for Innovation KNIME

Live Demo - Workflows

© 2020 KNIME AG. All rights reserved.

KNIME Hub

HUB.KNIME.COM

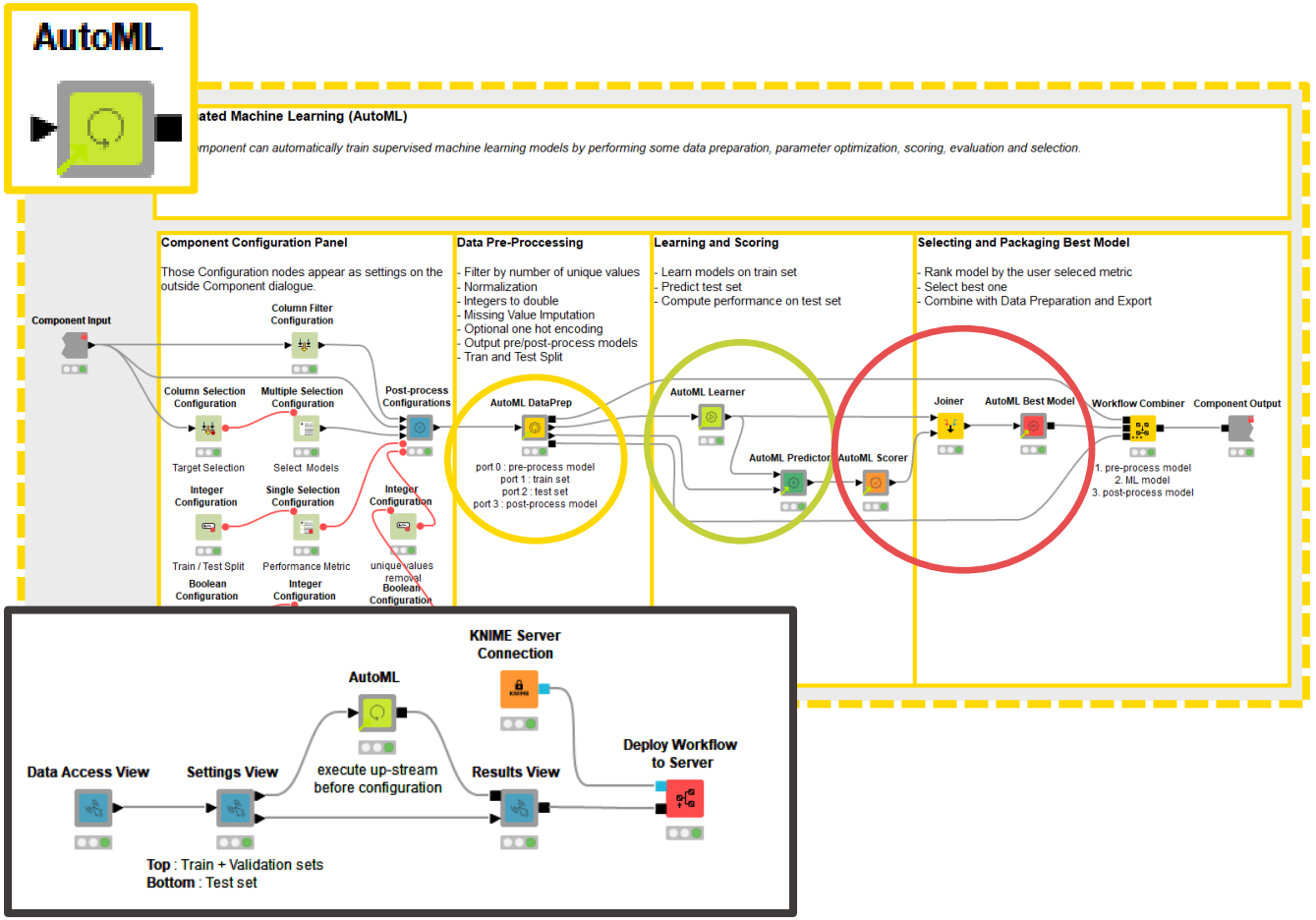


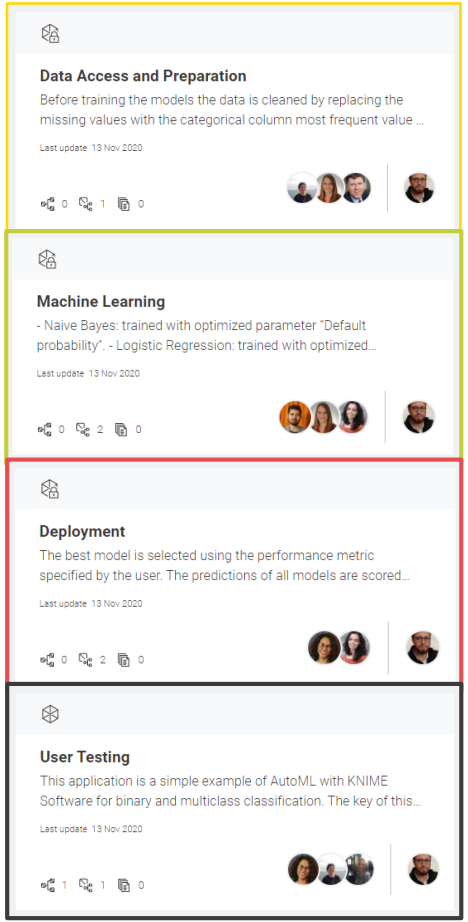
…. a peak into what's coming:

Using KNIME Hub capabilities in the Enterprise

© 2020 KNIME AG. All rights reserved.

Collaboration on a complex Component

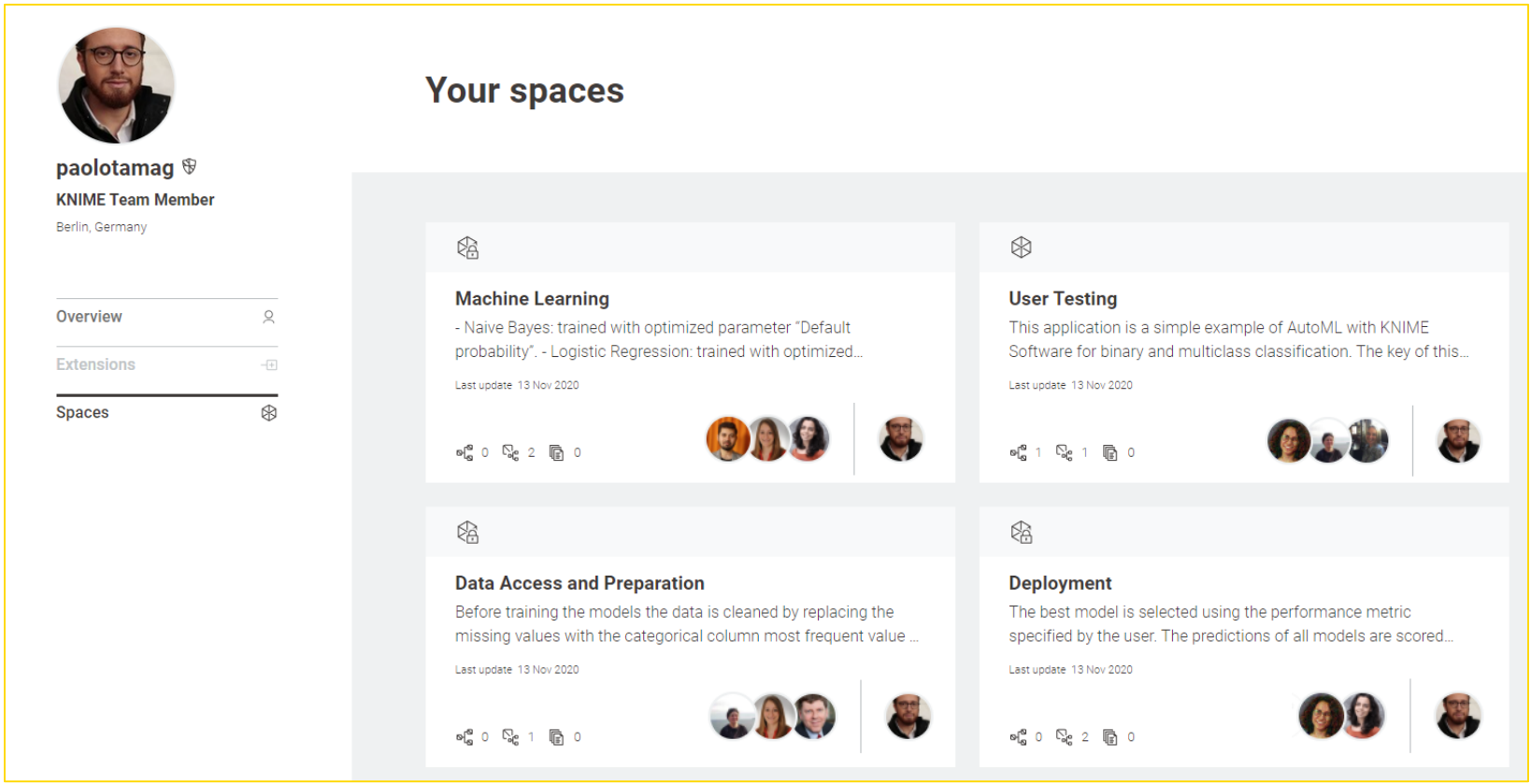




© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

Collaboration on our Component

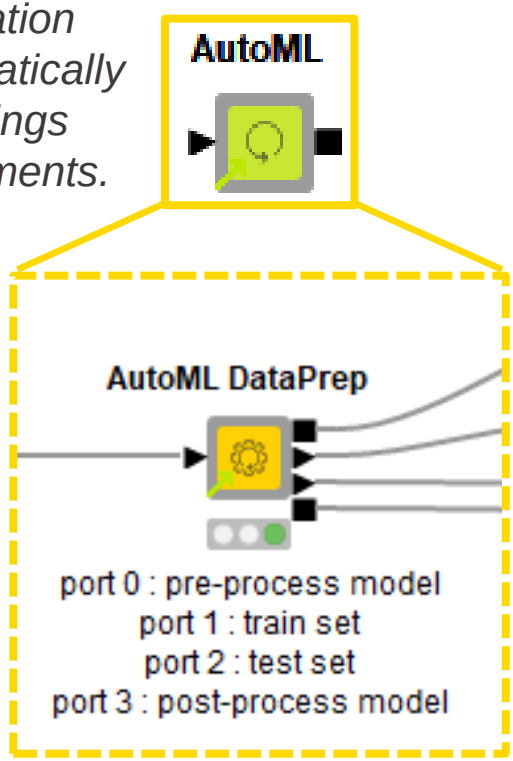


© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

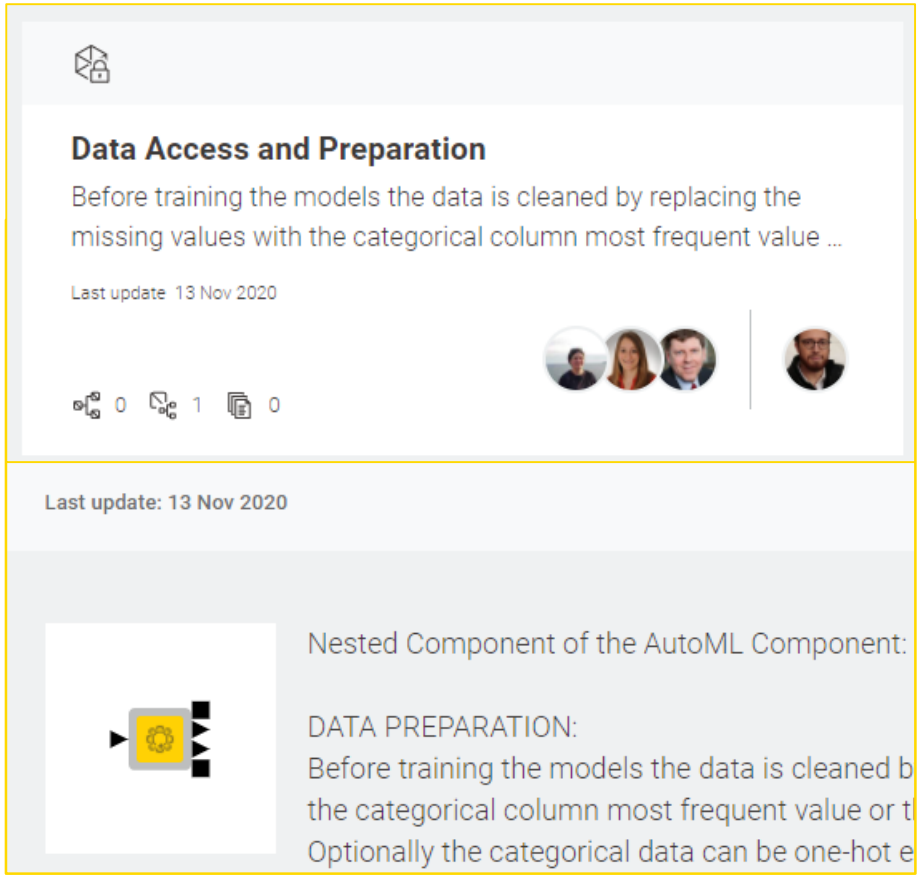
“Data Access and Preparation” Collaboration Space

A team making sure all the data preparation would place automatically based on user settings and model requirements.



Collaborators:

- Project Owner
- Data Engineer
- Data Scientist
- Business Analyst



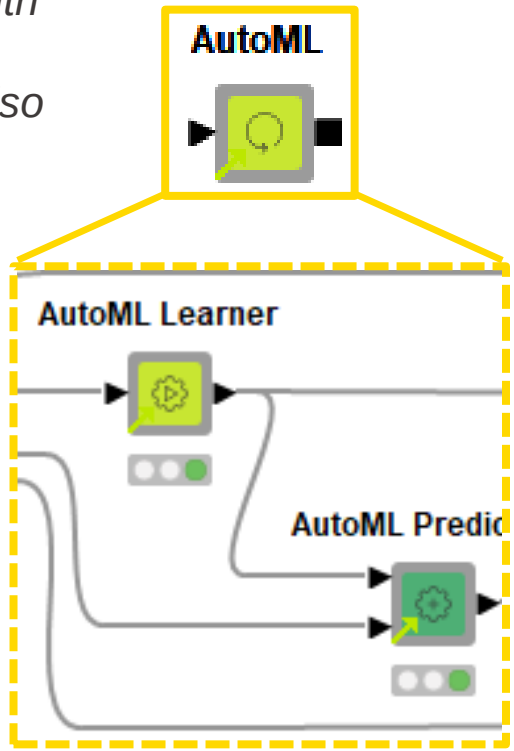
© 2020 KNIME AG. All rights reserved.

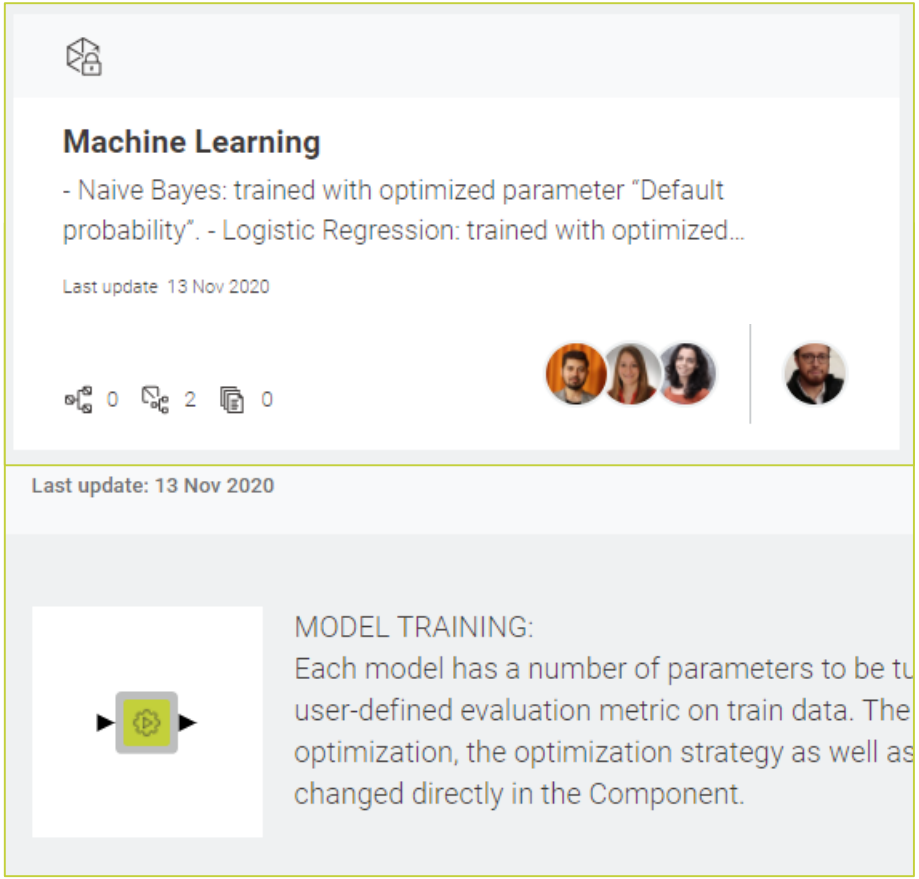
“Machine Learning” Collaboration Space

A team for training a variety of models with KNIME nodes and custom code, but also to be easily moved to production.

Collaborators:

- Project Owner
- Code First
- Data Scientist
- ML Ops





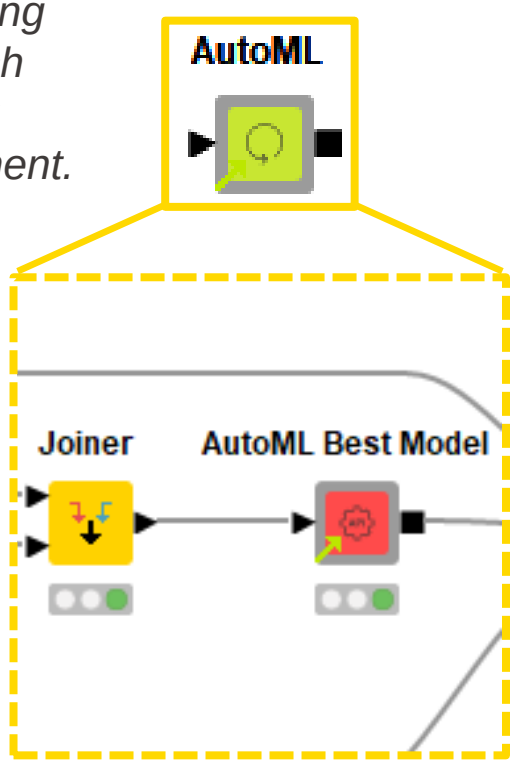
© 2020 KNIME AG. All rights reserved.

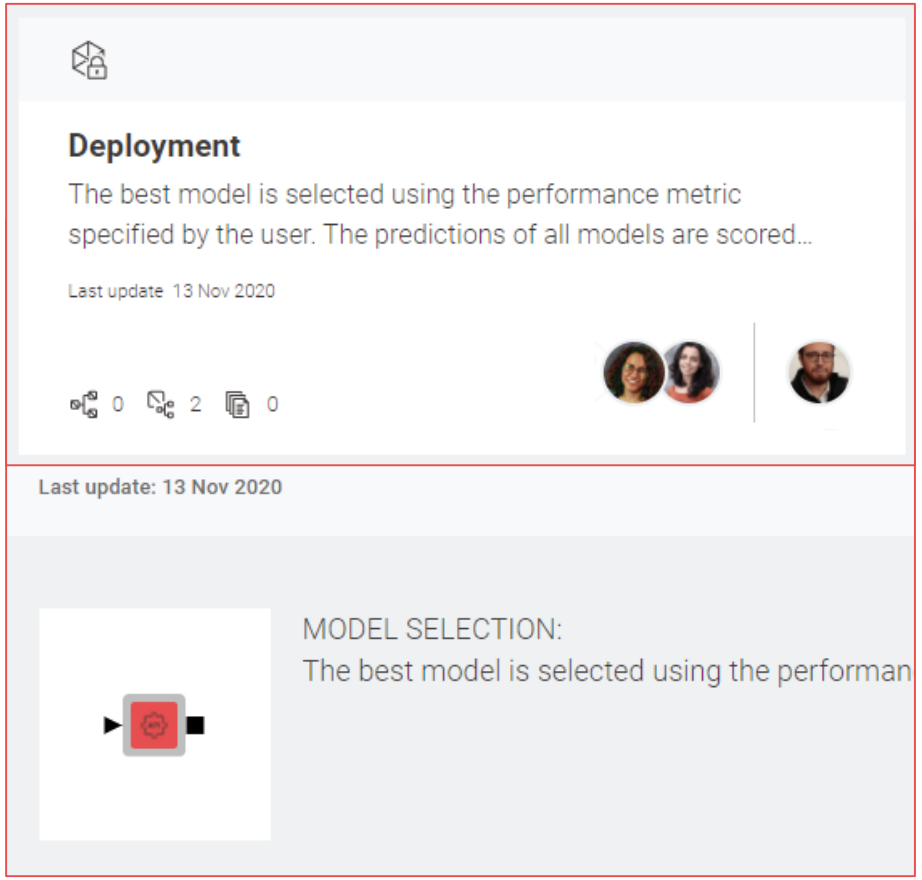
“Deployment” Collaboration Space

A team for selecting, building and exporting a single model which satisfies policies for the correct deployment.

Collaborators:

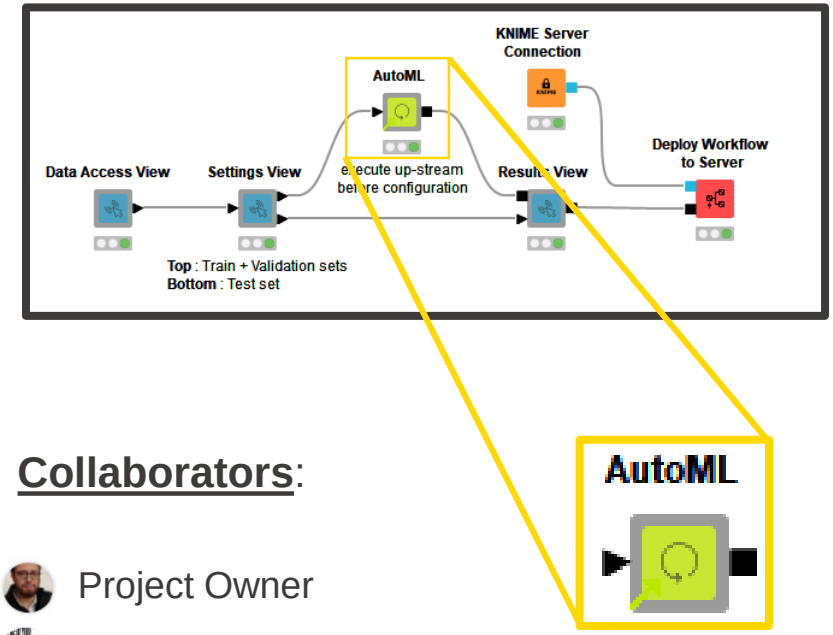
- Project Owner
- ML Ops
- DS Practice Leader





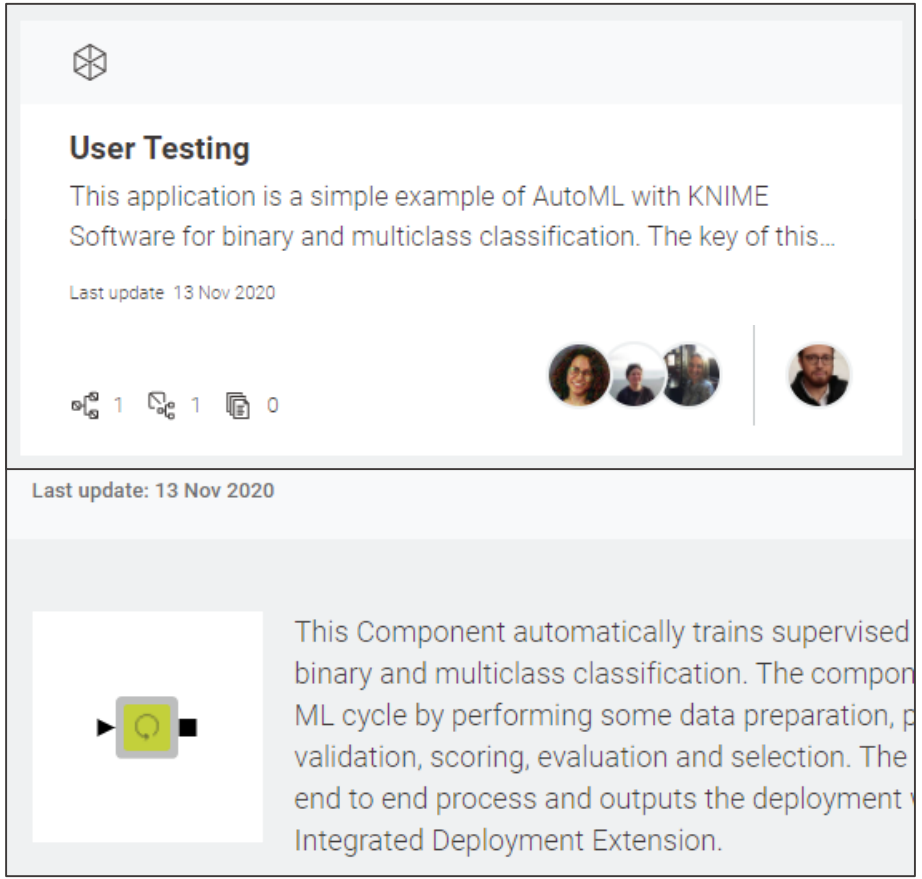
© 2020 KNIME AG. All rights reserved.

“User Testing” Collaboration Space



Collaborators:

- Project Owner
- Information Consumer
- Business Analyst
- DS Practice Leader

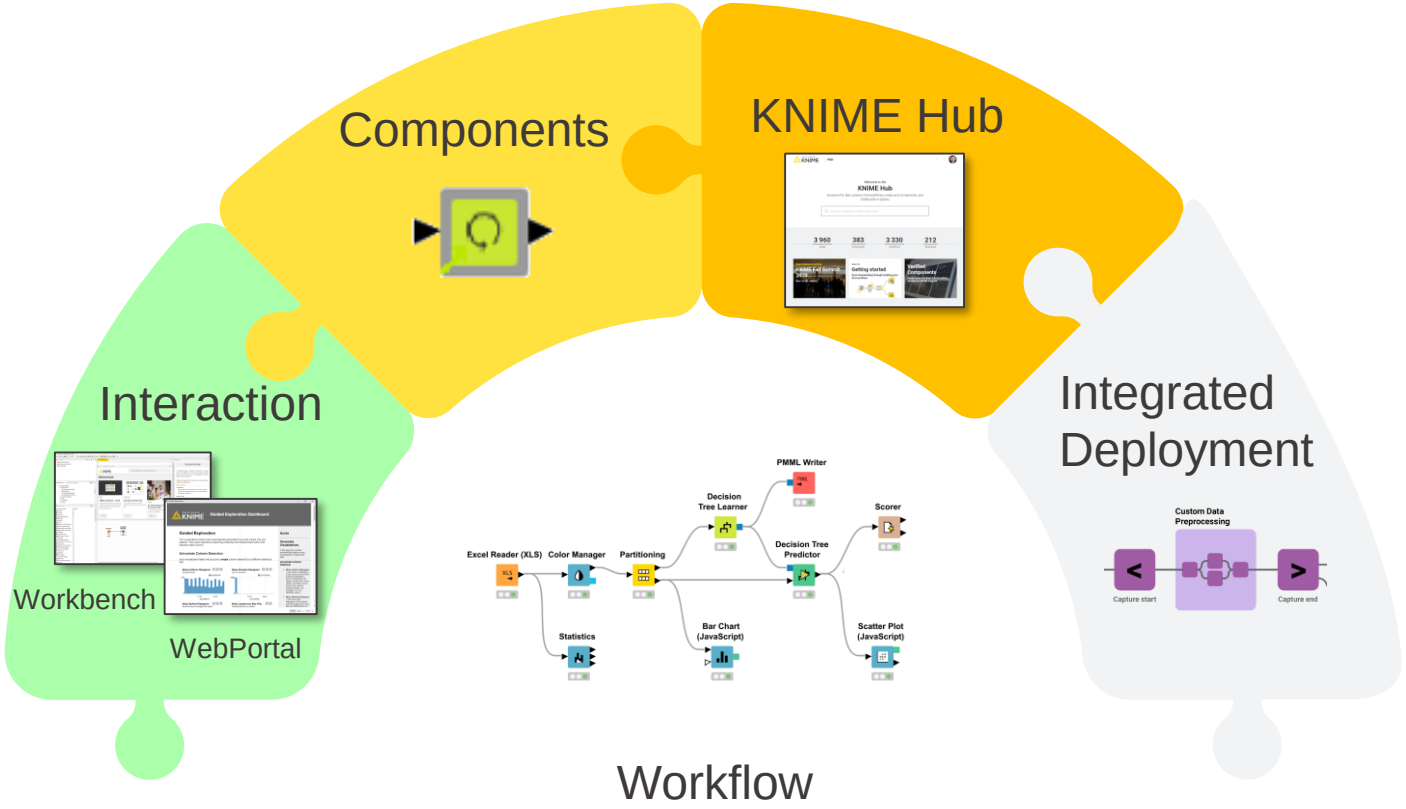


© 2020 KNIME AG. All rights reserved.

Live Demo – Collaboration

© 2020 KNIME AG. All rights reserved.

Highlights for Team Collaboration



© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

New KNIME Pros Learnathon Tomorrow!

Group 1
Financial Analysis

Maarit
KNIME Team Member

Lada
KNIME Team Member

Group 2
Life Sciences

temesgen-dadi
KNIME Team Member

Francosinus
KNIME Team Member

Group 3
Automation

paolotamag
KNIME Team Member

Mpattadkal
KNIME Team Member

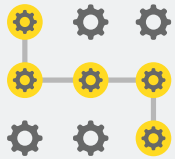
© 2020 KNIME AG. All rights reserved.

Successful Team Collaboration

Team Leader
Data Science Practice

Data Engineers
Data Scientists
“Code First” / Package First

Business Analyst
Interactive User
Information Consumer



Model / ML Operations

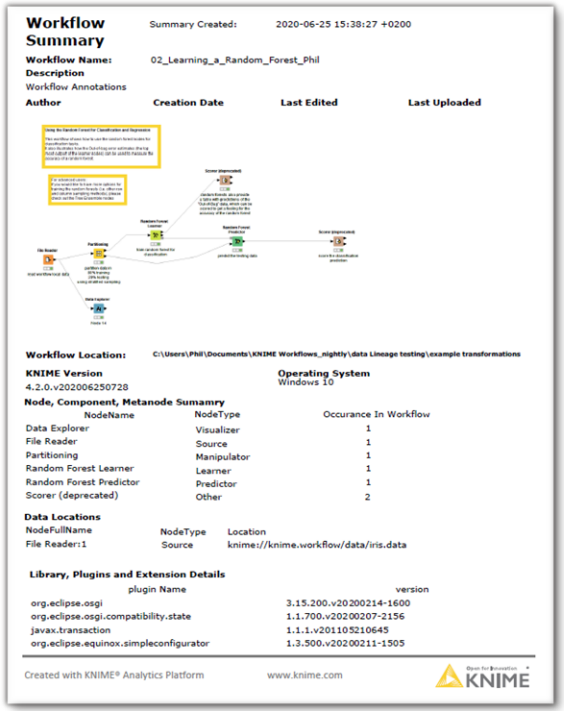
IT Operations
Scalability and control
Centralized resources / strategies
Security

Compliance
Financial / risk oversight
Costs allocation
Governance, traceability, GDPR

© 2020 KNIME AG. All rights reserved.

Compliance*: Governance, Auditing, GDPR (Best Practices!)

Metadata: External Documentation; Data / Model Lineage, Asset identification



Workflow Summary

Summary Created: 2020-06-25 15:38:27 +0200

Workflow Name: 02_Learning_a_Random_Forest_Phil

Description

Workflow Annotations

Author | **Creation Date** | **Last Edited** | **Last Uploaded**

Workflow Location: C:\Users\Phil\Documents\KNIME Workflows_nightly\data Lineage testing\example transformations

KNIME Version
4.2.0.v202006250728

Operating System
Windows 10

Node, Component, Metanode Sumamry

NodeName	NodeType	Occurance In Workflow
Data Explorer	Visualizer	1
File Reader	Source	1
Partitioning	Manipulator	1
Random Forest Learner	Learner	1
Random Forest Predictor	Predictor	1
Scorer (deprecated)	Other	2

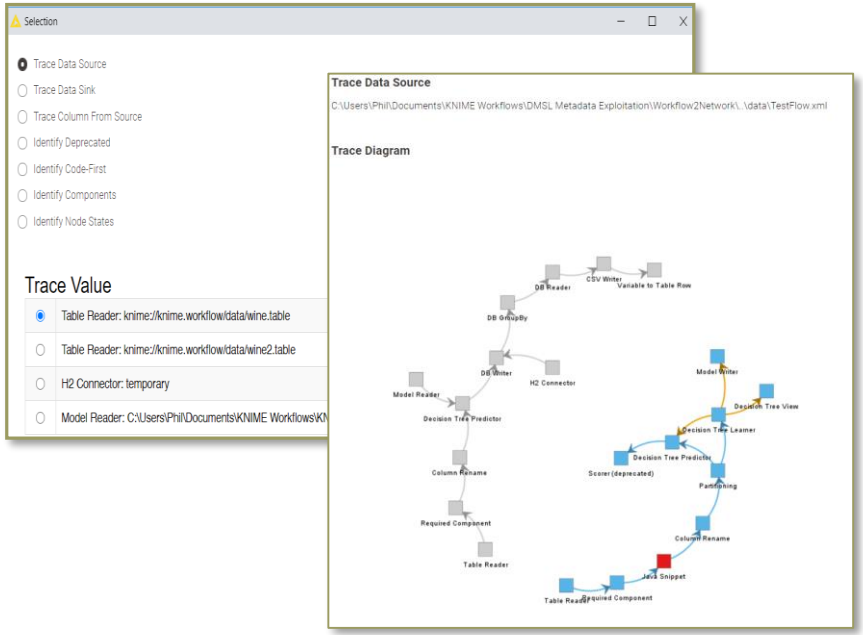
Data Locations

NodeFullName	NodeType	Location
File Reader:1	Source	knime://knime.workflow/data/iris.data

Library, Plugins and Extension Details

plugin Name	version
org.eclipse.osgi	3.15.200.v20200214-1600
org.eclipse.osgi.compatibility.state	1.1.700.v20200207-2156
javax.transaction	1.1.1.v201105210645
org.eclipse.equinox.simpleconfigurator	1.3.500.v20200211-1505

Created with KNIME® Analytics Platform www.knime.com KNIME

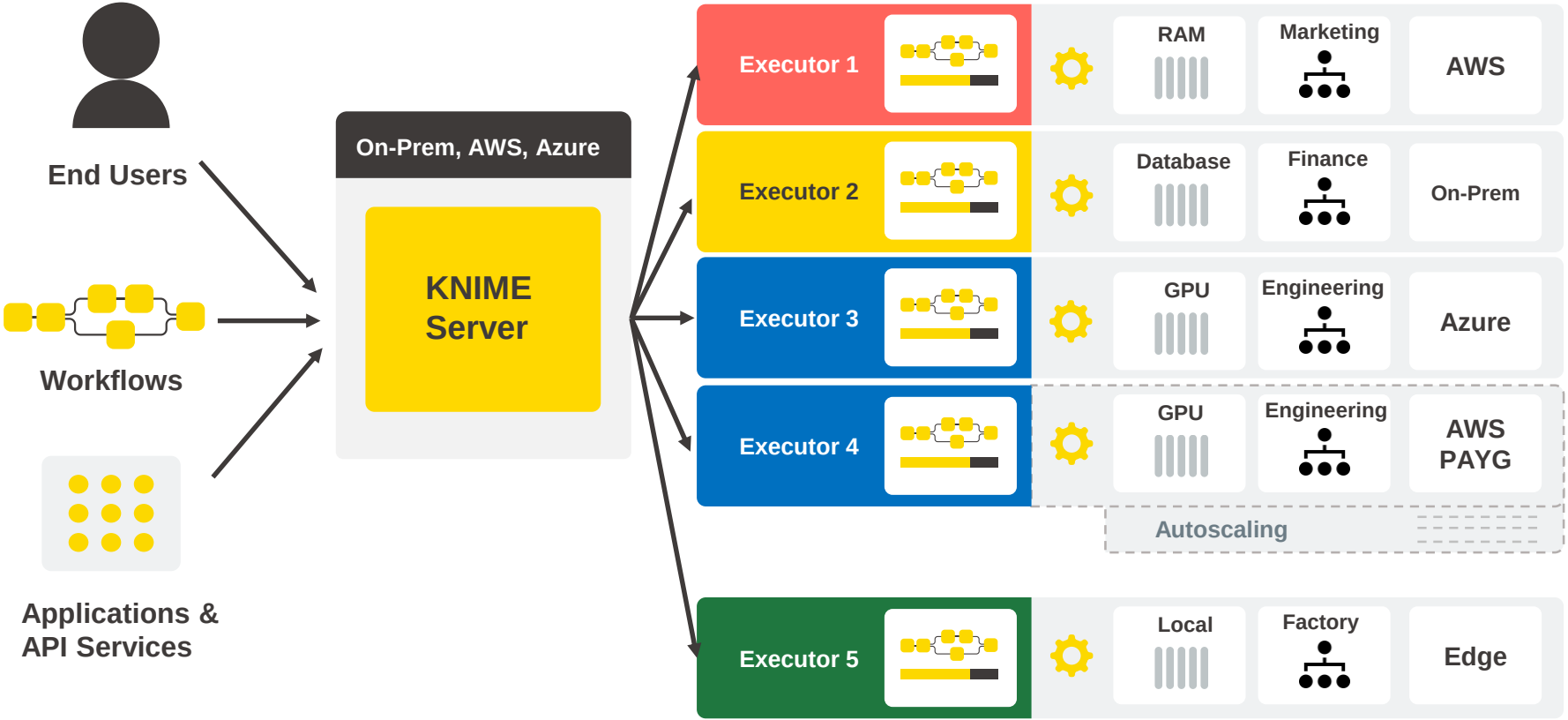


*Don't tell them, but compliance folks are just Information Consumers / Interaction Users!

© 2020 KNIME AG. All rights reserved.

Open for Innovation KNIME

IT Operations: Choice, control and definable costs



© 2020 KNIME AG. All rights reserved.

Successful Team Collaboration

Team Leader
Data Science Practice

Data Engineers
Data Scientists
“Code First”
/ Package First

Open for Innovation
KNIME

Business Analyst
Interactive User
Information Consumer

Model / ML Operations

IT Operations
Scalability and control
Centralized resources / strategies
Security

Compliance
Financial / risk oversight
Costs allocation
Governance, traceability, GDPR

© 2020 KNIME AG. All rights reserved.

THE END

© 2020 KNIME AG. All rights reserved.