Open for Innovation ®

KNIME
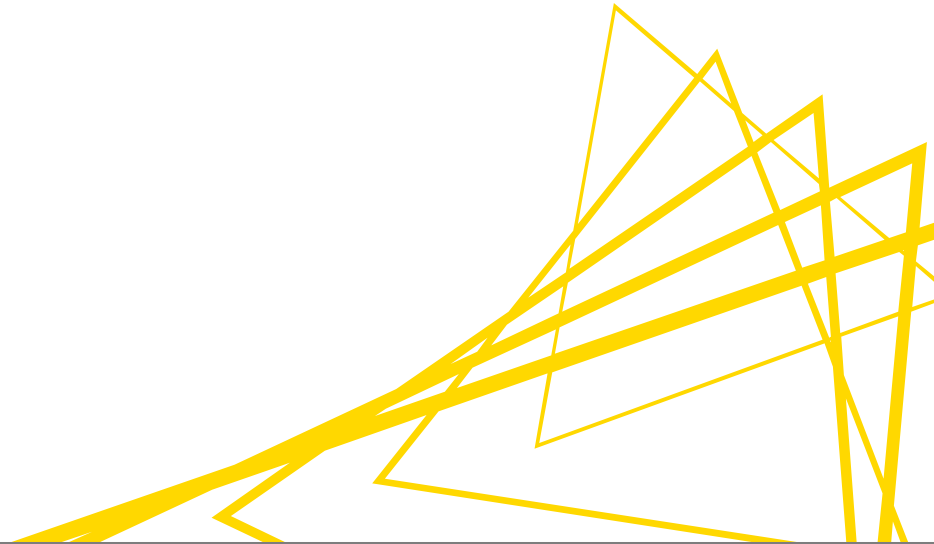
KNIME Fall Summit 2018
- Opening -

Jim Falgout & Michael Berthold

KNIME

#KNIMESummit2018

The Plan...

- Recent KNIME highlights

- Hypes and Trends: AI, ML, DS,... Automation!

- The Summit.

Open for Innovation   ®
KNIME

A (small) Subset of Highlights

Internal Highlights: New (US) KNIMErs

- Jim Falgout
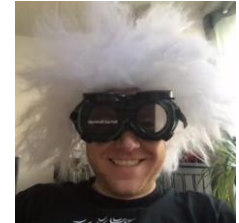- Jeff Gullick

- Scott Fincher

- Cynthia Padilla

- Jason Tyler
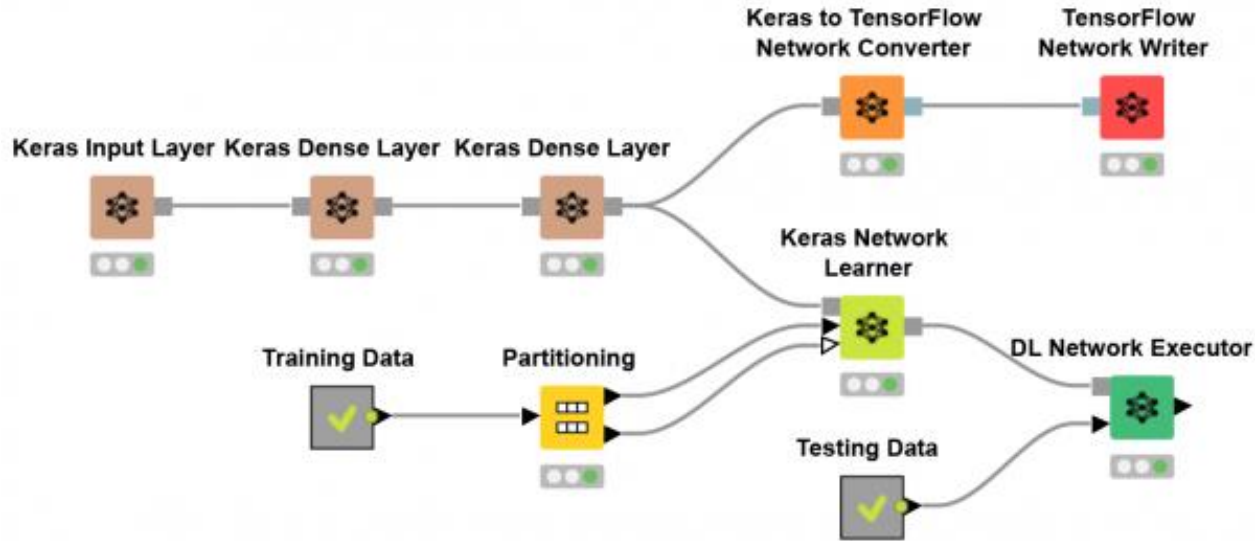
- David Butler

- Paul Treichler

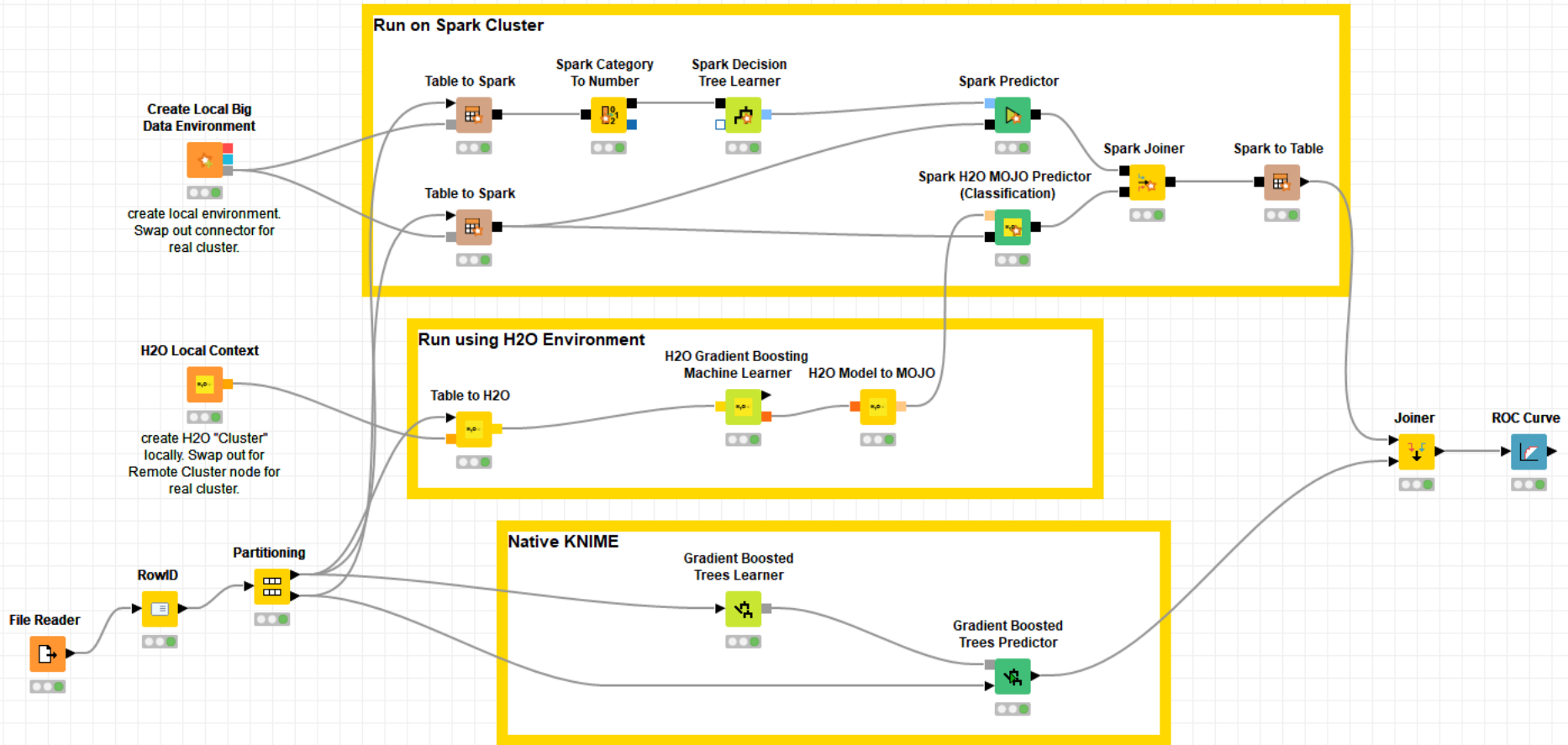- Jason Denzin

One EU Highlight:

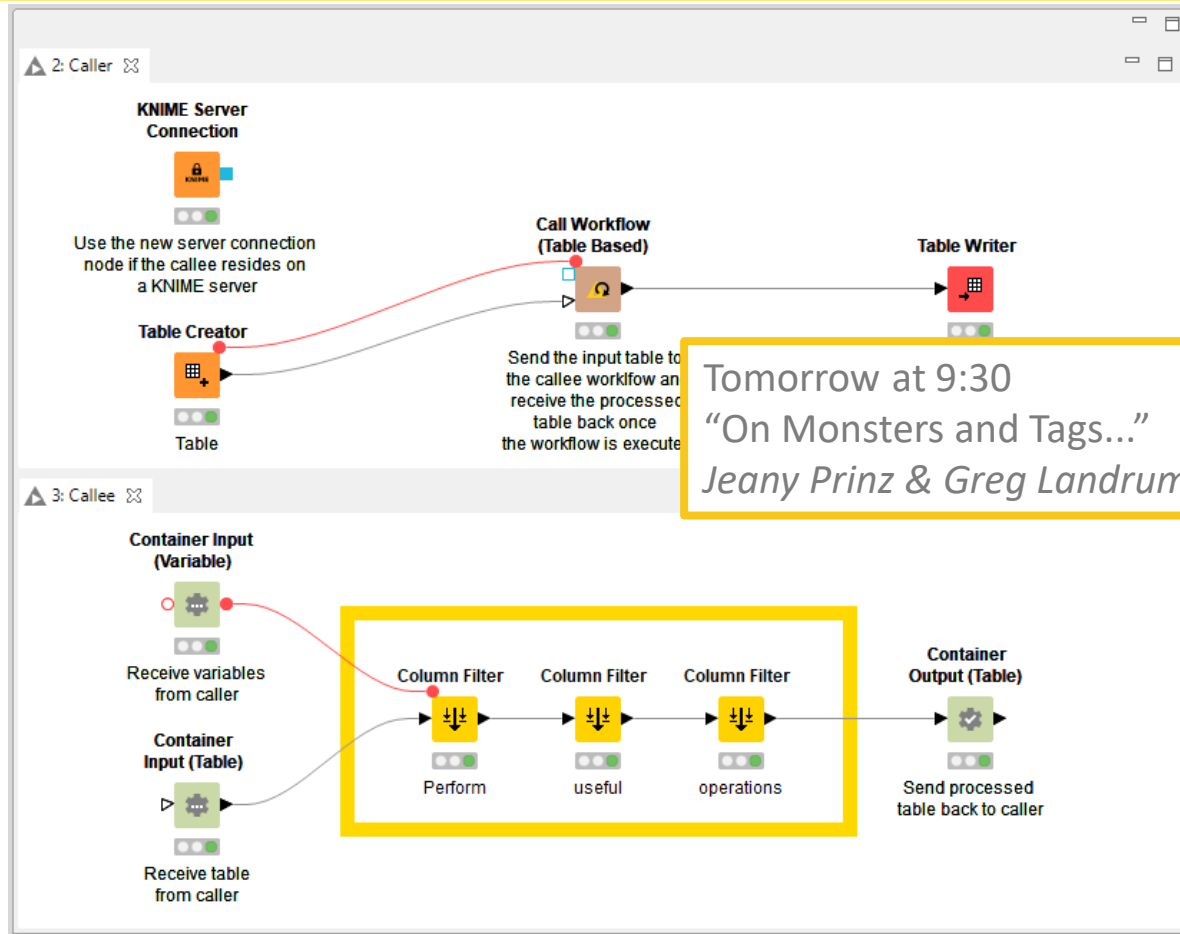- Christiane Kallfaß

A few Tech Highlights: Tool Blending

What's New Session, Various mentions, and tomorrow at 10:00
"…KNIME… for High-Throughput Image Analysis"
Andries Zjiestra (Vanderbilt/Nashville)

Open for Innovation ®
KNIME

Talking about Blending…

A few Tech Highlights: Automation & Management

Two Tech Highlights for Dean…
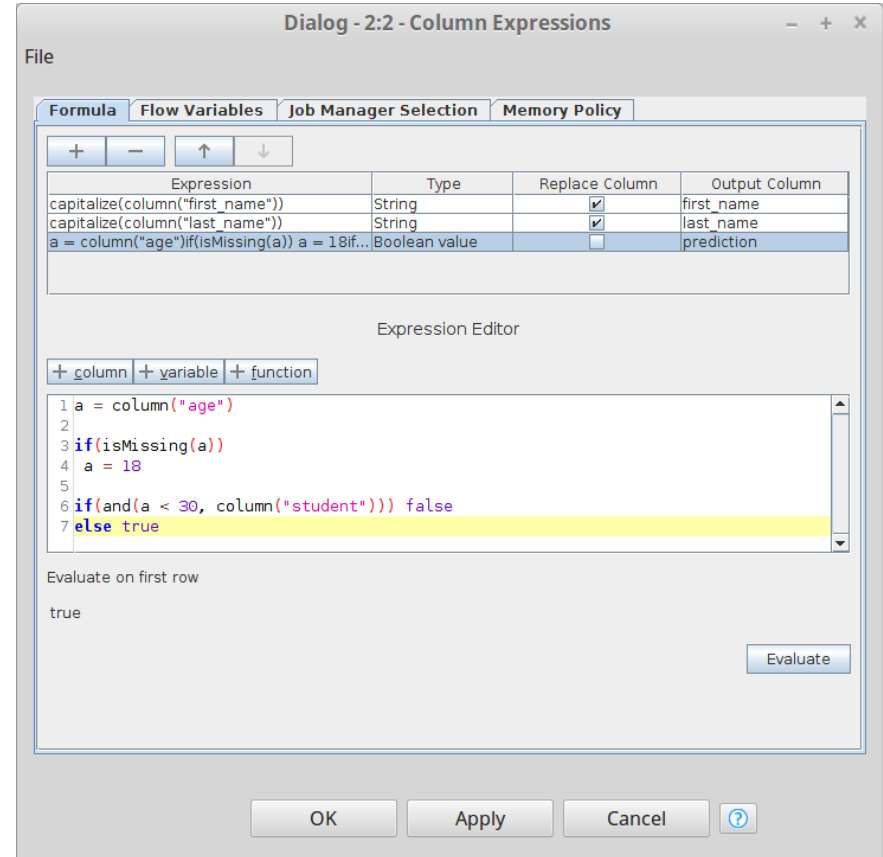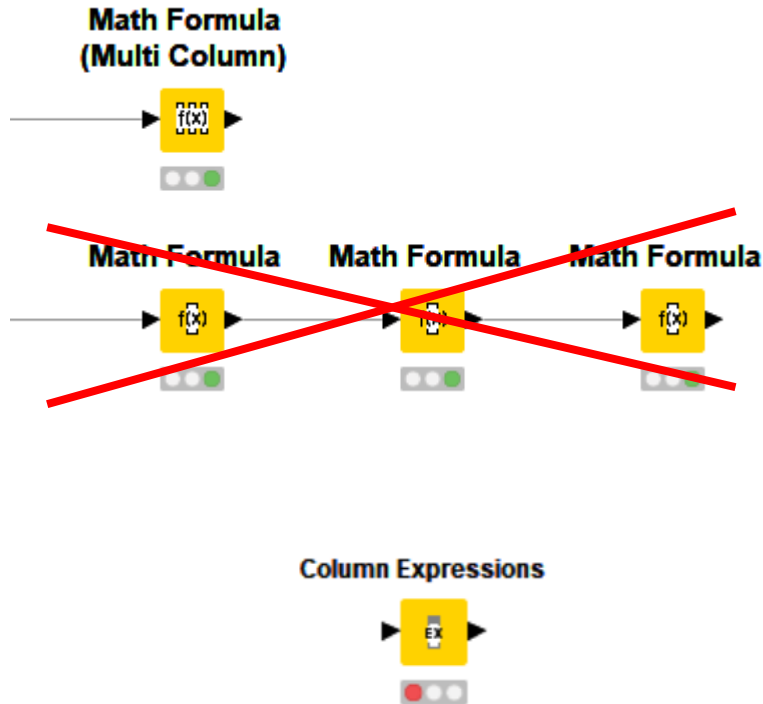
Tech Highlights

many more...:

- New Data Types & Sources

- More Tools in the Blender

- Beautiful Visualizations

- Performance and Management on the Server

- Containerization of Workflows

→ What's New Session!

Open for Innovation ®
KNIME

A Side Note on Software Quality

WILEY — **WIREs** DATA MINING AND KNOWLEDGE DISCOVERY

FOCUS ARTICLE

On the accuracy of linear regression routines in some data mining packages

B. D.

¹Departm
LeBow C
Philadelp

²Departm
Florida, C

Correspo
B. D. Mc
Sciences
Drexel Un
Email: bd

Conclusions

[…]

1. Of the five that offer the sample variance, ▊ employs a bad algorithm.
2. All seven offer linear regression, and **only IBM and KNIME offer reliable algorithms**.
3. Of the two that offer ANOVA, ▊ has a bad algorithm.
4. The ▊ polynomial regression procedure appears to be extremely inaccurate.

[…]

Open for Innovation ®
KNIME

Let's talk about Trends!

Technology Trends

- Artificial Intelligence, Machine Learning, Deep Learning, Data Science, …

- Automation (Driverless, Robotic, Magic…) AI/ML

- Productionizing Data Science

Artificial Deep Intelligent What?

- Machine Learning
  - Learn a Model to fit some data
    (predict airline delays, adjust heat, control valve, …)
  - Or: improve performance of an agent from feedback
- Artificial Intelligence
  - Learn to Perform a Human-like Task
    (recognize objects in images, words in speech signals, …)
  - **Pragmatic** AI often mistaken for **Pure** AI*

Mike Gualtieri: "Artificial Intelligence: Fact, Fiction")

ZDNet 🔍   **WINDOWS 10**   **CLOUD**   **INNOVATION**   **SECURITY**   **TECH PRO**   **MORE** ▾   **NEWSLETTERS**   **ALL WRITERS**

📄 MUST READ: Apple Mac mini (2018) review: The little Mac that could

Google ponders the shortcomings of machine learning

Scientists of AI at Google's Google Brain and DeepMind units acknowledge machine learning is falling short of human cognition and propose that using models of networks might be a way to find relations between things that allow computers to generalize more broadly about the world.

By Tiernan Ray | October 20, 2018 -- 12:52 GMT (05:52 PDT) | Topic: Artificial Intelligence

Open for Innovation ®
KNIME

Artificial Deep Intelligent What?

- Machine Learning   **→ Methods!**
  - Learn a Model to fit some data
    (predict airline delays, adjust heat, control valve, …)
  - Or: improve performance of an agent from feedback

- Artificial Intelligence   **→ Tasks!**
  - Learn to Perform a Human-like Task
    (recognize objects in images, words in speech signals, …)
  - **Pragmatic** AI often mistaken for **Pure** AI*

- Data Analysis, (Advanced/Visual) Analytics…   **→ Insights!**
  - Extract (understandable!) Knowledge from Data
    (customer behavior, reason for churn, …)

Mike Gualtieri: "Artificial Intelligence: Fact, Fiction")

Open for Innovation ®
KNIME

ML vs AI vs Data Science

Machine Learning:
Train Models for
- Predictions / Forecasts
- Clusters / Groupings
- Anomalies / Outliers
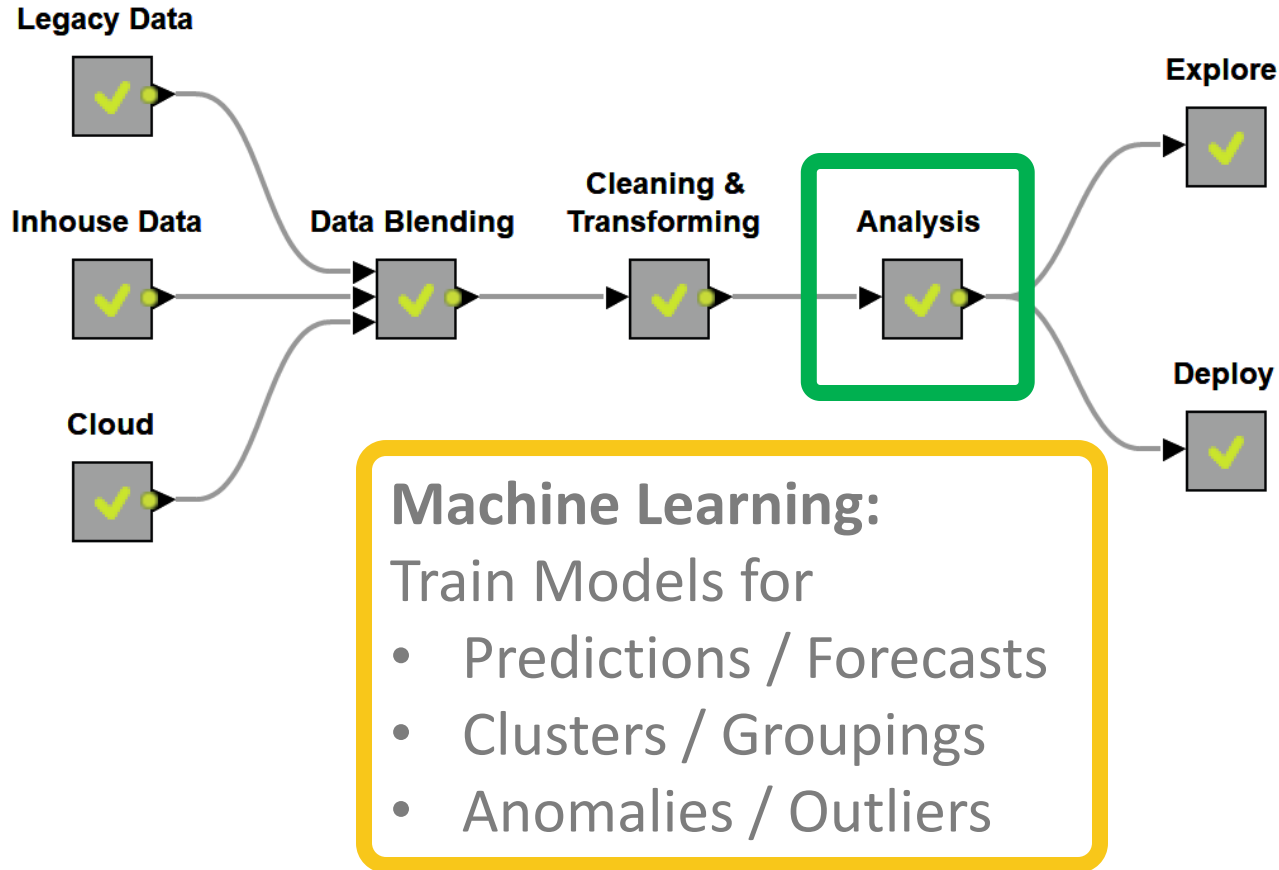
Machine Learning and Deep Learning

- Classic Machine Learning (incl. Neural Networks):
  - assume data is ready to go

  > *Scott Fahlman:*
  > *"if you have great preprocessing, you don't need a learner"*

- Deep Learning
  - More Power = Learn More from More Data
  - Move Learning also towards Data Preparation

- Learning works best with background knowledge!

J.S. Denker(?): "Neural Networks are the second best way to do almost anything"

Open for Innovation ®
KNIME

ML vs AI vs Data Science

Deep Learning: Train Models that also transform data

ML vs AI vs Data Science

Legacy Data

Inhous

Cloud

Today at 5pm
"Doing the Data Science Dance:
The Interplay Between Algorithms, Data Preparation, Sampling, and Interpretation"
Dean Abbott

Deploy

Data Science:

Applications that
- create Models
- provide Summaries
- trigger Insights
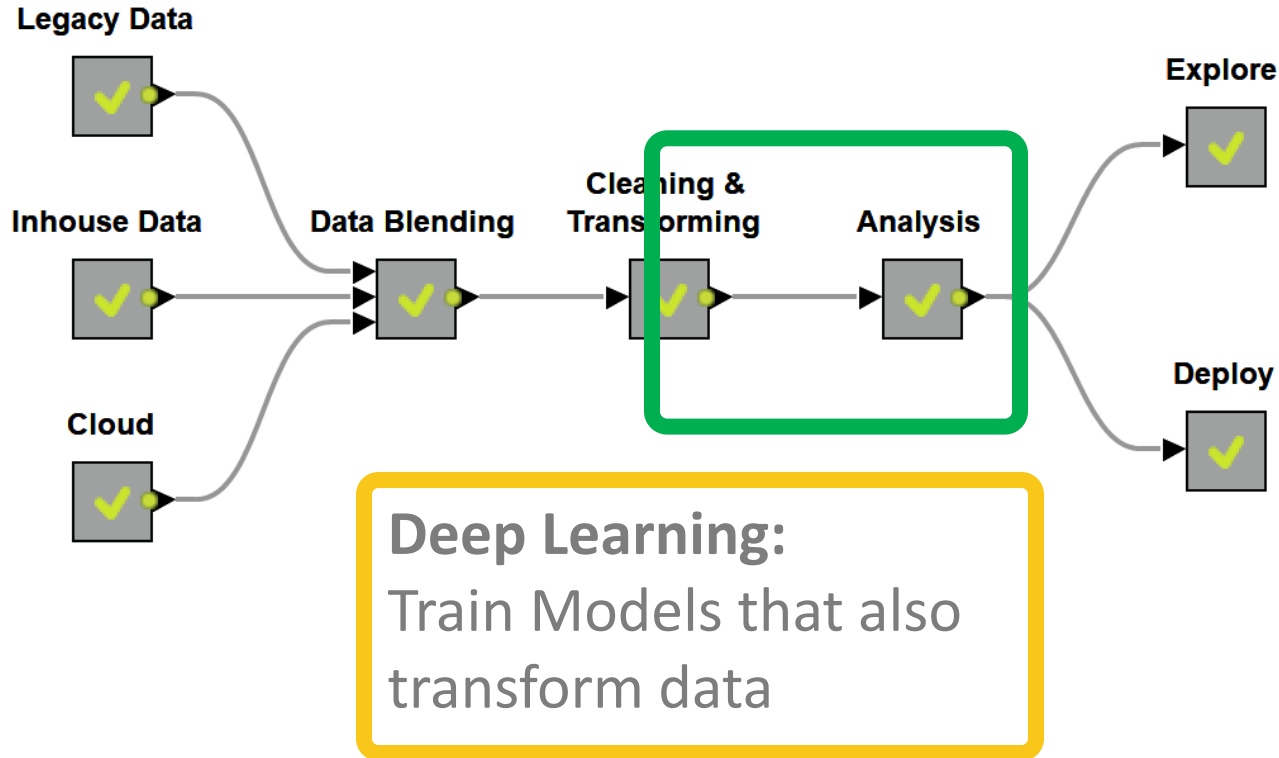
Open for Innovation ®
KNIME
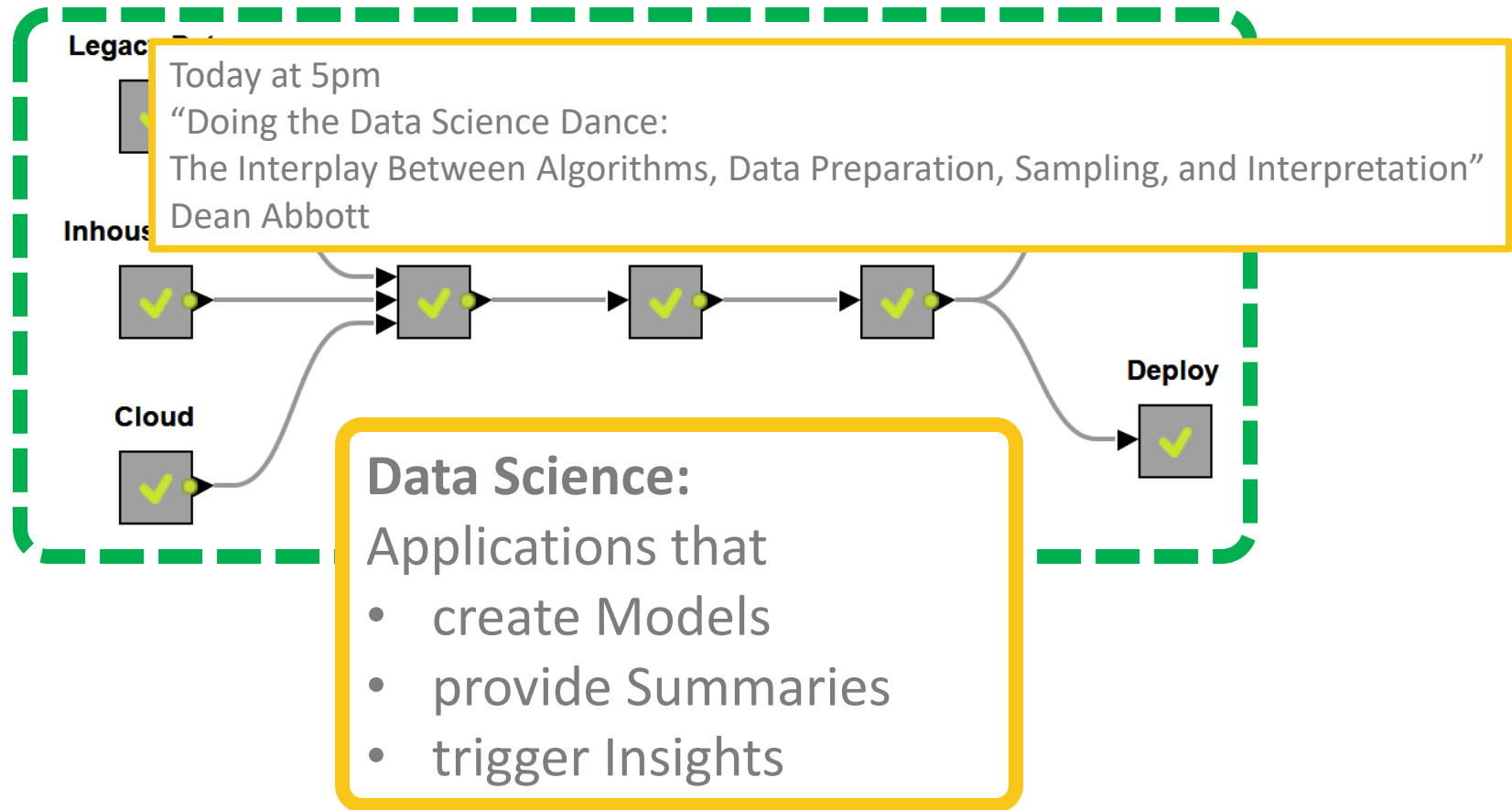
Automating What?

Automating Data Integration:
- Parsing
- Record Matching
- ...
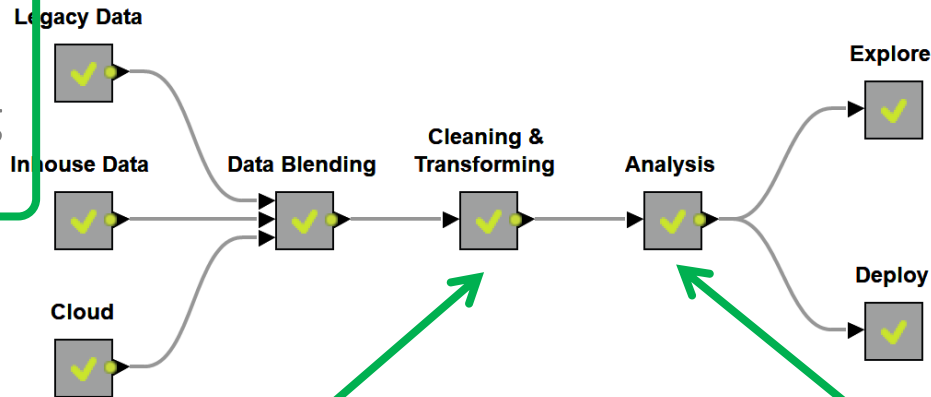
Automating Data Proc:
- Feature Selection
- Feature Construction
- Data Cleaning

Automating Analytics:
- Parameter Optimization
- Model Selection
- Ensemble Construction

Automating How?

- Try: (Brute Force)

  – Use compute power to try variations

- Copy: (Transfer Learning)

  – Use similar problems as starting point

- Learn: (Bayesian Optimization)

  – Use Machine Learning to predict/learn parameter space

Open for Innovation ®

KNIME

Automating Everything?

Human Input Needed!
- Data Selection – Is this relevant?!
- Analysis Goal – What is interesting?
- Exploration – This looks weird?...

Automating Data Integration:
- Parsing
- Record Matching
- ...

Legacy Data

In-house Data

Data Blending

Cleaning & Transforming

Analysis

Explore

Cloud

Deploy

Automating Data Proc:
- Feature Selection
- Feature Construction
- Data Cleaning

Automating Analytics:
- Parameter Optimization
- Model Selection
- Ensemble Construction

Open for Innovation ®
KNIME

Automating How?

- Try: (Brute Force)

  – Use compute power to try variations

- Copy: (Transfer Learning)

  – Use similar problems as starting point

- Learn: (Bayesian Optimization)

  – Use Machine Learning to predict/learn parameter space

- Interact: (Guided Automation)

  – Augment Automation with <u>Human</u> Intelligence

Open for Innovation ®

KNIME

Flexible Automation / Interaction

- Data Scientists build (or start from template)
  - automate the boring pieces
  - add interaction where human feedback is needed
- Deploy to Business Users as Analytical App
  - hides complexity
  - enables interaction at the right level of detail

(How much interaction? That depends…)

Open for Innovation   ®
KNIME

KNIME's Guided Automation: Automation + Interaction

Interaction Points

www.myserver.ch

Tomorrow at 2pm
"Guided Analytics for ML/AI Automation"
Christian Dietz & Simon Schmid (KNIME)
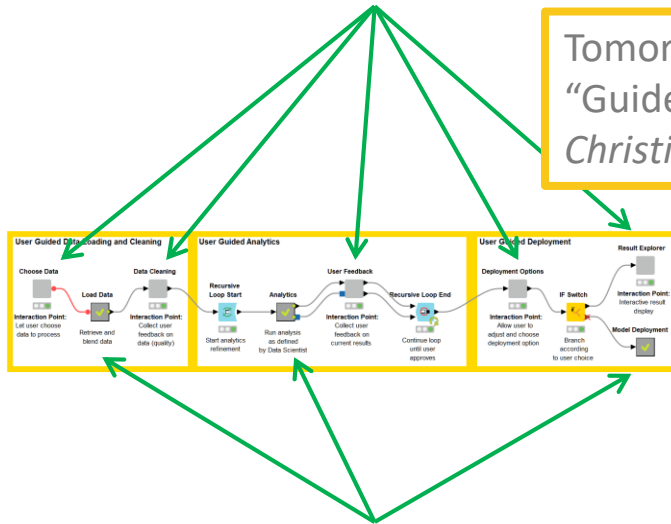
Automated
(prepackaged, updated, or created by in-house experts)

KNIME Server

26

Open for Innovation ®
KNIME
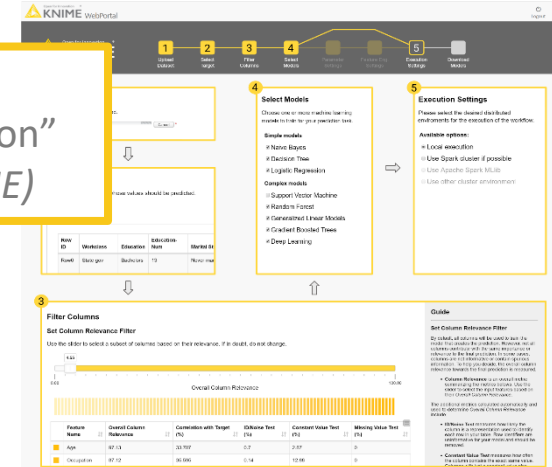
What's Next?

- AI Hype will start to cool down
  - some applications successful ("pragmatic" AI)
  - getting closer to real cognition ("pure" AI) out of reach
  - next step: make AI/ML part of the mix
- Automation will hype a bit more
  - potential for automation of well defined problems
  - move to production a lot harder (remember Big Data?)
  - flexible platforms increasingly important (productionizing Data Science)

→ Augment human expertise, don't replace it!

KNIME Fall Summit 2018

Open for Innovation ®

KNIME

KNIME Spring Summit 2018

THURSDAY, NOVEMBER 8

9:00 AM		
	REGISTRATION, COFFEE, BREAKFAST	
10:00 AM		**Welcome & Opening** Michael Berthold (KNIME)
10:30 AM	**KNIME SESSION**	**What's New & Cooking** Bernd Wiswedel & Team (KNIME)
1:00 PM		
	LUNCH BREAK IN TEJAS DINING ROOM	
2:00 PM		**Finding Themes in Text Data to Help Transform Member Experience** Melvi M. Methippara (Kaiser Permanente)
2:30 PM	**SESSION 1**	**Using Analytics to Improve Consumer Choice in the US & the UK** Michelle Leonard, Doris Sullivan (Consumer Reports)
3:00 PM		**Advanced Job Analytics @ Daimler** Julian Leweling (Daimler)
3:30 PM		**COFFEE BREAK**

Open for Innovation ®
KNIME

3:30 PM		**COFFEE BREAK**
4:00 PM		**Data Science at Palo Alto Networks: How Do We Innovate?**
		Nandan Thor & Sirish Upadhyay (Palo Alto Networks)
4:30 PM	**SESSION 2**	**Data Science at Palo Alto Networks: How Do We Productionize?**
		Juho Parviainen & Nilesh Dhomse (Palo Alto Networks)
5:00 PM		**Keynote Presentation**
		Dean Abbott (Smarter HQ)
6:00 PM		**END OF SUMMIT DAY ONE**
6:40 PM - 7:05 PM	**BUSES TO MICHELADAS**	Take the elevator from M1 to Level M2 and make a U-turn to the right. Go out the doors to the new Rowling Hall Plaza. Buses will travel from Whitis Avenue between 6:40 PM and 7:05 PM. **Bring your KNIME Fall Summit name tag to identify yourself for the bus ride and dinner.**
7:00 PM		**DINNER AT MICHELADAS**
8:30 PM - 11:00 PM	**SHUTTLE BUS TO AT&T**	From 8:30 PM, a bus will shuttle between Micheladas and the AT&T. The last bus will leave Micheladas at 11:00 PM.

Open for Innovation ®
KNIME

3:30 PM

4:00 PM

4:30 PM

5:00 PM

6:00 PM

6:40 PM
-
7:05 PM   **BUSES TO MICHELADAS**

7:00 PM

8:30 PM
-
11:00 PM   **SHUTTLE BUS TO AT&T**

FRIDAY, NOVEMBER 9

8:30am:

REGISTRATION, COFFEE, BREAKFAST

9am: Short Server Intro
Jon Fuller (KNIME)

9:30 AM		On Monsters and Tags...
		Jeany Prinz & Greg Landrum (KNIME)
10:00 AM	**SESSION 3**	**Deploying KNIME in an Amazon Cloud Environment**
		for High-Throughput Image Analysis
		Andries Zjiestra (Vanderbilt/Nashville)
10:30 AM		**A Data Pipeline Approach to Orphan Disease Insights**
		Sebastian Lefebvre (Alexion Pharmaceuticals)

11:00 AM

COFFEE BREAK

11:30 AM		**Guided Analytics at Seagate**
		Allan Luk & Eric Lin (Seagate)
12:00 PM	**SESSION 4**	**Data Analytics in Data Storage Device Development & Testing**
		Debin Wang (Seagate)
12:30 PM		**Using KNIME for Optimizing Die Utilization**
		Zachary Eich (AMD)

Open for Innovation ®
KNIME

1:00 PM		**LUNCH BREAK IN TEJAS DINING ROOM**
2:00 PM		**Guided Analytics for ML/AI Automation**
		Christian Dietz & Simon Schmid (KNIME)
2:30 PM	SESSION 5	**Enterprise Scale Data Blending**
		Shalini Subramanian (Juniper Networks)
2:50 PM		**REST API: Workflow Integration with Python**
		Owen Watson (Juniper Networks)
3:15 PM		**COFFEE BREAK**
3:30 PM		**Custom Language Translation using KNIME & Keras**
		Mohammed Ayub & Joseph Gochal (NFPA)
4:00 PM	SESSION 6	**Creating an Equipment Anomaly Detection Framework**
		Ziad Katrib (Calpine)
4:30 PM		**Turning AI Hype into Something Practical: Demystifying Bots**
		Phil Winters & Vincenzo Tursi (KNIME)
5:00 PM		**FAREWELL RECEPTION/OPEN BAR**

Open for Innovation ®
KNIME