Opening

Jim Falgout  
(GM, KNIME US)

Michael Berthold  
(CEO, KNIME)

#KNIMESummit2019
Agenda

• End to End Data Science
  – What does that mean in Real Life?
  – How does KNIME Software fit in
  – What about Automation?

• A Few 2019 Highlights

• The 2019 Fall Summit
“End to End” Data Science: More Than Just Methods?
Data Science: More Than Just Methods!

KNIME Analytics Platform

KNIME Extensions  KNIME Integrations  Community Extensions  Partner Extensions

KNIME Server

KNIME WebPortal  Data Science as a Service
KNIME Software for the Entire End to End Data Science Cycle
Gather & Wrangle

- A (hidden) Strength
- Hundreds of Connector & Transformation Nodes
- Big Data Extensions
- New: Extended Cloud File System Connectivity
Model & Visualize

- State of the Art Methods: Hundreds of Native Nodes
- Integrations of Best of Breed Libraries and Tools
  - Application Specific (Life Science, Text Processing, Image & Time Series Analysis, ...)
  - Specialized Implementations (H2O, XGBoost, ...)
  - Coding Support (but not a requirement!) - (R, Python, ...)
- Continued Integration of Deep Learning Environments
- New: Cloud Cognitive Services
Deploy & Manage

• Deployment to Humans & Machines
  – Analytics Applications
  – Web Services

• New
  – Modularization of KNIME WebPortal
  – Distributed & Pinned Execution
  – Simplified Packaging of Services (Container I/O)
Deploy & Manage

The KNIME Model Factory
• Retrain always, on demand, when needed
• Automated Model Initialization
• Monitoring, Alerting, ...

Production Support / Governance
• Versioning
• Backwards Compatibility
• Documentation
• Compliance / Best Practices
• ...

White Paper
The KNIME Model Factory
Scaling Modeling Processes for the Enterprise

https://www.knime.com/blog/the-knime-model-process-factory
Consume & Optimize

• Consumption through direct Deployment:
  – Workflow as a Service
  – Workflow as an Analytics Application

• Optimization = User Feedback Incorporation
  – “Agile” Data Science: It’s all in the Workflow
  – Built-In Optimization: (Inter-) Active Learning
Automation?

Gather & Wrangle → Model & Visualize → Deploy & Manage → Consume & Optimize
1\textsuperscript{st} Generation Automation

1. Clean Data In

2. Pre-canned Interaction

3. Model Out

Gather & Wrangle

Start & Manage

Consume & Optimize
2\textsuperscript{nd} Generation Automation: Automate & Guide

1. Read Data In
2. Customized Interactions
3. Interactive Use / Continuous Deployment

Gather & Wrangle
Model & Visualize
Deploy & Manage
Consume & Optimize
2nd Generation Automation: Automate & Guide

1. Read Data In
2. Customized Interactions
3. Interactive Use / Continuous Deployment

Gather & Wrangle

Model & Visualize

Deploy & Manage

Consume & Optimize
2\textsuperscript{nd} Generation Automation: Mix & Match

Automated

Customized

Guided Interaction

Gather & Wrangle

Deploy & Manage

Consume & Optimize
Abstracting Data Science?
Data Science Abstraction
Data Science Abstraction

Applications

Services

Building Blocks

Workflows

Languages
Components in KNIME:

- Continued Abstraction of Data Science: Automated or Guided...
  - ...Machine Learning
  - ...Feature Selection & Engineering
  - ...Interpretation
  - ...Model Management & Deployment
  - ...

- Sharing of Reusable Pieces
  - Integrating & Blending Corporate Data
  - Standardize Data Access
  - Define Corporate Reporting Schema
  - Gray Boxes (want to know/change what’s inside?)
  - Best Practices, Compliance, ...
  - ...
Which Type of Data Science?

Standard Problems without Business Impact
(Others can do it better or have the data)
No in-house Domain or Data Science Expertise Required
(One (data) model fits it all – & for a while...)
Productionizing Data Science is Critical for Value Creation
(Continuous incorporation of new tech & user feedback)
Data Science Innovation Deeply Embedded in Business
(Exploring latest ML/AI trends, injecting own R&D into DS)

⇒ Standard Tools / Services
⇒ Automation / Outsourcing
⇒ Custom Data Science
⇒ Mix Custom DS with Code
Data Science in Production needs Agile Environments

Agile Data Science in a Nutshell:
• Define Initial Requirements (1. Goal, 2. Data...)
• Develop Initial Analytical Application/Service
• Deploy to Business Users
• Gather Feedback
• Adjust/Expand Requirements
• Add Interaction, Automation,...
• (Try out weird new Algorithms)
• Go Back (in short cycles)

⇒ Requires Real End to End Environment
   (One Platform for Entire Cycle & all Levels of Abstraction)
2019 Highlights
KNIME Continues to Grow

• The Company
  – 100+ KNIMErs
  – KNIME US tripled
  – New Leaders:
    • Albert Gashi (Revenue)
    • Tilman Eberle (Marketing)
  – Just joined: Lindsay
  – Just signed: Jip, Stephen

• The Community
  – More than 250,000 active
  – More than 50,000 trained
  – Over 100 Meetup groups world wide
Welcome to the
KNIME Hub

The place to find and collaborate on KNIME workflows and nodes. Here you can find solutions for your data science questions.

Search workflows, nodes and more...

<table>
<thead>
<tr>
<th>Nodes</th>
<th>Components</th>
<th>Workflows</th>
<th>Extensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,542</td>
<td>169</td>
<td>2,116</td>
<td>180</td>
</tr>
</tbody>
</table>
### Thursday, Part 1

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 AM</td>
<td>Registration, coffee, breakfast</td>
</tr>
<tr>
<td>10:00 AM</td>
<td><strong>Opening Session</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Opening Comments</strong></td>
</tr>
<tr>
<td></td>
<td>Michael Berthold (KNIME)</td>
</tr>
<tr>
<td></td>
<td><strong>Evangelism, Education, Partner, and Community News</strong></td>
</tr>
<tr>
<td></td>
<td>The KNIME Team</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Coffee break</td>
</tr>
<tr>
<td>12:00 PM</td>
<td><strong>Session 1: Data Science in Action</strong></td>
</tr>
<tr>
<td></td>
<td>Running a 24x7 Mission Critical Data Platform for the World's #1 Media Agency</td>
</tr>
<tr>
<td></td>
<td>Laura McElhinney (Horizon Media)</td>
</tr>
<tr>
<td></td>
<td><strong>Operating at Scale: Building a Self-Service Community</strong></td>
</tr>
<tr>
<td></td>
<td>Robert Fritts (AT&amp;T)</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>Lunch break - Tejas Dining Room</td>
</tr>
<tr>
<td>2:00 PM</td>
<td><strong>Session 2: KNIME in the Enterprise</strong></td>
</tr>
<tr>
<td></td>
<td>What's New and Cooking in KNIME Server</td>
</tr>
<tr>
<td></td>
<td>Jim Falgout (KNIME)</td>
</tr>
<tr>
<td></td>
<td><strong>Moving Decisions to the Edge with Guided Analytics at Seagate</strong></td>
</tr>
<tr>
<td></td>
<td>Allan Luk, Brendan Doherty, Debin Wang (Seagate Technology)</td>
</tr>
<tr>
<td></td>
<td><strong>KNIME and Mathematics Keeps the Optician Away</strong></td>
</tr>
<tr>
<td></td>
<td>Hakan Genc and Sharon Boczkiewicz (Siemens Digital Industries Software)</td>
</tr>
</tbody>
</table>

Last year’s monster session now spread out...
Thursday, Part 2

3:30 PM  Coffee break

4:00 PM  Session 3: Components are for Sharing
What’s New and Cooking in KNIME Analytics Platform
Bernd Wiswedel (KNIME)

Auditing Human Decisions in the Background Check Industry
SJ Porter (HireRight)

Visualizing with Guided Analytics: For End Users AND Experts
Phil Winters and Team (KNIME)

5:45 PM  End of day

6:15 PM &
6:30 PM  Shuttle buses to dinner

6:30 PM  Dinner - Banger's Sausage House & Beer Garden

Bus 1 - 6:15 PM  • Buses 1 & 2 depart from AT&T at these times. Meet at the Plaza Rowling Hall Exit.
Bus 2 - 6:30 PM

6:30 PM - 11:15 PM  • Join us for BBQ, beers, and greens at Banger’s Sausage House & Beer Garden.
• A bus will shuttle between Banger’s and AT&T approximately every 15 minutes between 6:30 PM and 11:15 PM.

CRAFT BEER
With 200+ beers on draft and another 50 in cans and bottles, we’ve got the largest draft system in the state of Texas.
### Friday

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 AM</td>
<td>Registration, coffee, breakfast</td>
</tr>
</tbody>
</table>
| 9:00 AM | **Session 4: Life and Other Data Science**  
**What's New and Cooking in the KNIME WebPortal**  
Jeany Prinz and Paolo Tamagnini (KNIME)  
**Building a Clinically Significant Rare Disease Data Master: Approach and Workflows**  
Sebastien Lefebvre (Alexion Pharmaceuticals)  
**Using KNIME to Build a Data-Driven Culture (and Workflows!) in a Biopharma Setting**  
Kenneth Longo (Wave Life Sciences) |
| 10:30 AM| Coffee break |
| 10:45 AM| **Session 5: KNIME in Government and Finance**  
*Virtual Anticipation Network: Applying Tensor Completion to the Study of Inter-State Competition and Irregular Conflict*  
Curtis Fox (Combating Terrorism Technical Support Office) and Brian Frutchey (NuWave Solutions)  
**Financial Applications of Self-Service ETL and Geospatial Analysis**  
Paras Gupta (BGIS)  
**Keynote Presentation**  
**Focus on the Task: Enabling your Models to Speak the Language of the Business**  
Dean Abbott (SmarterHQ) |
| 1:00 PM | Lunch break - Tejas Dining Room |
| 2:00 PM | End of KNIME Fall Summit 2019 |

Upon rfr*: stopping with lunch this year.

(*) aka: reasonably frequent request
A Call for KNIME Champions

We would love to understand:
• How you’re using KNIME Software
• How long you’ve been using it
• Your project highlights

And if you’d be interested in:
• Testing new features?
• Being a reference?
• Writing about your project highlights?
• Giving a presentation?

Yes?
Come to the KNIME Champions table in the coffee break area. We’re next to the KNIME Demo Booth

Any movie stars out there? Let us know if we could film a short interview with you!