Opening Remarks
KNIME Spring Summit 2019

Michael Berthold

20 March 2019

#KNIMESummit2019
Predicting Attendance...

Number of Attendees: **378 (March, 15th)**; **409 (today)**

Jan 15: 103 -> prediction: 379
March 7: 333 -> prediction: 377
Predicting Attendance: The Dirty Truth...
Outline Analysis

• Diagnostic    -  2018 (and one early 2019) Highlight
• Predictive    -  Trends (+ a tad Prescriptive Analytics)
• Descriptive   -  Spring Summit Revisited
Diagnostic:
Recent Highlights
Other Highlights

• Lots of new cool Technology
• New Users, Customers, and Partners – 24 talks & two parallel tracks this week!

• Greg’s Highlight: “Vintage Website hauled into 21st century”
• Office Expansions (Konstanz*2, Berlin*3, Austin...?)
• Community growth...
Community Growth on the KNIME Forum

Martin Lauber
Ivan Pazin
Igor Zaychik

Philip Katz
Armin Grudd
Steve Roughley
Community Blog Contributions

Srinivas Attili
A Forest of Tools and Islands of Information – A Data Science Journey
Author: Srinivas Attili, Director – Marketing Analytics & Data Science at Juniper Networks
Often, organizations adopt tools and systems as per the emerging technology trends and their organic needs. This leads to a situation where a company’s technology landscape often resembles a forest of tools crowded in one place, however each tool is different from the next.
Information created by these diverse tools is often known only to a small group of people within the company and remains isolated from such pools created by various other systems. These unconnected islands of information need to be brought together, synthesized to derive insights that would not have been realized with the siloes of data.
Read more

Ted Hartnell
Market Simulation with KNIME: Android vs iOS
Author: Ted Hartnell (CTO of Scientific Strategy)
What is Market Simulation?
A market simulation is a way to model a real world market. Just as real world markets have products, features, brands, stores, locations, and competitive rivals, so does a market simulation. But what makes a market simulation truly realistic are the customers. Simulations can generate tens of thousands of virtual customers designed to mimic the purchase decisions of real world shoppers. Customers evaluate the differentiation offered by each product.
Read more

Alexander Franke
Ever sat next to a friend or colleague at the computer and were awed when you see certain tasks is much better? We recently asked KNIME users to share their series of posts we’ll be showing you how the experts use KNIME in the hopes that by sharing ideas you’ll discover some handy techniques.
So where do bunny ears come into this?

Chris Baddeley
Metanodes for Reusability: A Short story of metanodes, wrapped metanodes, and metanode templates.
Authors: Chris Baddeley and Rosanne Stefo
What is a Metanode?
Before we start, what is a metanode? Metanodes are gray nodes that contain sub-workflows. They play the role of functions or macros in script based tools.
They look like a single node, although they can contain many nodes and even more metanodes.
Read more

Craig Cullum
Snowflake meets Tableau, Data warehouse in an hour?
In this blog series we’ll be experimenting with the most interesting blends of data and tools. Whether it’s mixing traditional sources with modern data lakes, open source savings on the cloud with protected internal legacy tools, SQL with NoSQL, web wisdom of the crowd with in house handwritten notes, or IoT sensor data with idle chatting, we’re curious to find out if they will blend? Want to find out what happens when IBM Watson meets Google News, Hadoophive meets Excel, R meets Python, or MS Word meets MongoDB?
Read more

Stephen Roughley
Kees Schippers
Danyal Khaliq
Vijay Venkataraman
Tips & Tricks for Using KNIME Analytics Platform.
Ever sat next to a friend or colleague at the computer and were awed when you see certain tasks is much better? We recently asked KNIME users to share their series of posts we’ll be showing you how the experts use KNIME in the hopes that by sharing ideas you’ll discover some handy techniques.

Viewing Content in the Workflow Editor
Community Groups around the World
Predictive: Trends
Trend of Trends

• Model Management (2017)
  – “The KNIME Model Factory” (Berlin 2017)

• Automation & Privacy (2018)
  – Guided Analytics (& Automation) (ongoing since 2015)
  – and various sessions today!

• Custom vs. Pre-Built/Automated Data Science (2019)
Recap - Data Science Personae

• Researchers: Inventing new Methods
  – KNIME Community Extensions

• Coders: Exploring bleeding-edge Analytics
  – R, Python, Deep Learning, ... Integrations

• Data Scientists: Blending Data, Tools & Expertise
  – KNIME Workflows for Collaboration

• Casual Users: Reusing Analytical Blueprints
  – KNIME Server for Sharing Templates (Workflows & Metanodes)

• Business Users: Using Analytical Applications & Services
  – KNIME Server (Containers & REST) & WebPortal (Applications)
2015: Making Everybody a Data Scientist

• Researchers
• Coders
• Data Scientists
• Casual Users
• Business Users
Automating Data Science

- Researchers
- Coders
- Data Scientists
- Casual Users
- Business Users

Automation:
- Automate Model Selection, Generation (and Optimization)
- Hope for the Black Box to keep up with the state of the art
- Hope for the Data to be ready
Packaging Pre-Built Data Science

• Researchers
• Coders
• Data Scientists
• Casual Users
• Business Users

Data Science APIs:
• Pre-configured/-trained Models
• Hope for the API to match your Data
• Hope the Model does what’s needed
• Hope for regular Updates (and guaranteed Reproducibility?...)
Automation vs APIs

• Data Science for Everybody
  Data Science APIs
  – for extremely well defined standard problems
    (“Predict Event Attendance Participation from Past Numbers”)
  – don’t count on continuous, backwards compatible upgrades

• Data Science Automation
  – for well defined problems without a prepackaged solution
    (“Predict Developer Beer Consumption based on Release stress”)
  – don’t need cutting edge performance

The Answer to Everything: Shrink-Wrapping/Pre-Built Data Science?

⇒ Solves standard problems with standard methods for standard data.
Why Custom Data Science?

You have:
• Real Data (new fields, new sources, new types, messy)
• Inhouse Domain and Data Science Expertise (or great partners...)
• Maybe even Hybrid Setups (mix of legacy, on-prem, cluster, cloud(s), ...)

You need:
• Above-Average Performance
• Collaboration - Domain and Data Science Experts work together
• Reliable, reproducible Results
• Backwards Compatibility - What your Team built last year still works today (Best Practices, Blueprints, Reusable Components...)
Accelerating Custom Data Science

• Don’t force Data Scientists to program (Linux never “needed” a GUI either…)
  – but don’t lock out code, integrate it!
• Room to play
  – Allow to explore new trends
• Easy Access to all Data
  – Data Wrangling for Everybody
• Simple Operationalization
  – Move to production... ...in the same environment
Wait: No Automation and APIs?

- Automation does help to optimize selections
- Data Science APIs do help to reuse the Proven

The Power is in the Mix!

Custom Data Science
- Automates the Boring
- Incorporates Standards
- Allows Interaction to focus on the Interesting
The Elephant in the Room
The Elephant in the Room

What about Artificial Intelligence?
The Elephant in the Room

Artificial Deep Intelligent What?

• Machine Learning  ➔ Methods!
  – Train a Model to fit some data  
    (predict airline delays, adjust heat, control valve, …)

• Artificial Intelligence  ➔ Tasks!
  – Train a Model to perform a Human-like Task  
    (recognize objects in images, words in speech signals, …)  
  – **Pragmatic** AI often mistaken for **Pure** AI*

• Data Science, Advanced Analytics...  ➔ Insights & Systems!
  – Extract Actionable Knowledge or Models from Data  
    (customer behavior, reason for churn, …)

* Mike Gualtieri: “Artificial Intelligence: Fact, Fiction”
ML vs AI vs Data Science

Classic Machine Learning:
Train Models for
- Predictions /Forecasts
- Clusters /Groupings
- Anomalies /Outliers
ML vs AI vs Data Science

Deep Learning:
Train Models that also (partially) transform data

J.S. Denker: "Neural Networks are the second best way to do almost anything"
Data Science:
Applications that
• create useful Models
• provide Summaries
• trigger Insights
Not yet...

WHAT AI DOES IN ITS FREE TIME...

Creative Design: Casiana Rimbu
Descriptive:
Spring Summit - Revisited
Changes...

In the Past: Increasingly Diverse Audience
Now: Multiple Parallel Tracks
   Today: KNIME for the Business / Data Scientists
   Tomorrow: Life Science, IoT/Automation, Customer Intelligence

In the Past: Increasingly Gigantic KNIME Sessions
Now: Distributing the KNIME sessions
   What’s New / What’s Cooking together with User Talks

And finally: moving Phil where he belongs: to a theatre!

Let us know if it worked!

4:00 PM  SESSION 3: DEPLOYING DATA SCIENCE
   Room C 01
   What's New for Deployment
   Christian Dietz and Team (KNIME)

9:30 AM  SESSION 4: KNIME ANALYTICS PLATFORM
   Room C 01
   What's New in KNIME Analytics Platform
   Bernd Wiswedel and Team (KNIME)

2:00 PM  SESSION 6: KNIME SERVER IN ACTION
   Room C 01
   What's New in KNIME Server
   Jim Falgout (KNIME)

6:15 PM  WALK TO KINO INTERNATIONAL

7:00 PM  Visualizing with Guided Analytics: For End Users AND Experts
   Phil Winters and Team (KNIME)

8:00 PM  DINNER @ KINO INTERNATIONAL
Talk to Us!

• Feedback on cool KNIME Usage
  – for us to learn
  – maybe you want to write, talk, blog about it with us?...
  – Special Focus: **Your** take on Scalability

• Talk to us at the Champion table!

• ...and maybe join Phil & Iris on stage tomorrow...

4:30 PM
Room C 01

SESSION 7: KNIME - SCALE, DEPTH AND FOCUS
Scaling and Performance Tuning KNIME: A Guide
Phil Winters, Iris Adä (KNIME)
Thank You – and enjoy the 2019 KNIME Summit!

<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>OPENING SESSION</td>
</tr>
<tr>
<td>Room C 01</td>
<td>Opening Comments</td>
</tr>
<tr>
<td></td>
<td>Michael Berthold (KNIME)</td>
</tr>
<tr>
<td></td>
<td>Evangelism, Education &amp; Community News</td>
</tr>
<tr>
<td></td>
<td>The KNIME Team</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>COFFEE BREAK</td>
</tr>
</tbody>
</table>