KNIME Spring Summit 2018 - Opening -

Michael Berthold
KNIME

#KNIMEESummit2018
The Plan...

- A look backwards: 2017 Highlights
- A look forward: Trends
- The Summit.
2017 Highlights
A (small) Subset
...one more:

- **Tom Erickson** joins KNIME Board

**Chairman and CEO**

**Acquia**

Jan 2009 – Dec 2017 • 9 yrs
Greater Boston Area

Led Acquia from $0 to $175M in revenue, establishing the first business to offer open source as a service. Recognized as the fastest growing private software company in the USA by Deloitte in 2013.
Highlights: Technology

• Integrations!
  – Deep Learning (continued)
  – Google Sheets
  – Audio&Speech
  – H2O
  – ...

• Dean’s Node (Multi Col Math!)

• Completely revised Date&Time Support

• Many more JS Views

• Lots more improvements under the hood
  – Python integration (rev)
  – Logistic Regression rewrite
  – Apache Spark (ML) continued

• ...now >2000 nodes!
Highlights: Stuff

• US Office in Austin, Texas (Jim, Jeff, Scott, _, _, ...)  
• Open Source vs. Commercial Simplified:  
  – Open Source Analytics Platform: Data Science  
  – KNIME Server: Collaboration, Automation, Deployment  
• Quoting Gartner:  
  “[...] open-source is becoming more common throughout the data science market. It enables people to innovate collaboratively, each contributing their own perspectives [...].”
“...KNIME's unwavering commitment to open-source...”

“...[KNIME] demonstrates a deep understanding of the market, a robust product strategy and strength across all use cases...”

@KNIME gained the ability to run @H2O.ai algorithms, so these two may be viewed as complementary, not competitors #Ecosystem #OpenSource
Let’s talk about Trends!
Trends

- Deep Learning
  - Pragmatic AI often mistaken for Pure AI (Mike Gualtieri: "Artificial Intelligence: Fact, Fiction"

- Automation (as in “Model Factories”)

- Standardization (notably: Governing Privacy Rules)

- Auto, Driverless, Intelligent, Super... AI/ML/...
  - doing it all automagically...

Today at 14:00
“KNIME in Action: The Monster Model Factory”

Thursday at 9:30
“KNIME in Action: GDPR, Taking a Proactive Approach”
Phil Winters (KNIME)
Automating what, exactly?
CRISP-DM: Automation?

Potential for Automation
Automating ML/AI ≠ Automating Data Science

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

Automating Analytics:
• Parameter Optimization
• Model Selection
• Ensemble Construction
Issues with Automation

Data Scientist injects:
- smart optimization methods
- self-adjusting learning methods
- own expertise

Business Analyst:
- can not provide domain knowledge
- can not adjust choices (features, models, ...)
- ...it’s read only.
Issues with Automation

But I am buying some secret sauce, right?

Well, not really...

• Bayesian Optimization: smart parameter selection
• Active Learning: iterative model refinement
• Evolutionary Methods: feature construction
• other AI methods...

The value lies in how it is put together.
Can’t we build this Ourselves?

• Transparent:
  Our own data scientists can build or add to it

• Interactive:
  Our business users can inject their expertise
Guided Analytics: Automation + Interaction

Interaction Points

Automated
(prepackaged, updated, or created by in-house experts)

Side Note:
This is a classic Active Learning Loop!
Inside an Interaction Point
Guided Analytics and Automated AI, ML, ...

In a Nutshell:

• Automation can take...
• ...but then it also takes...

Guided Analytics allows:

• to automate the boring...
• ...but keep the expert in the loop.

Integration matters. Nobody has all the pieces.
On that “Integration” Note...
Partners at the Summit (talk, table, ...)
Partners at the Summit (attending)
**KNIME Spring Summit 2018**

**Wednesday March 7, 09:00 – 17:15, Room Berlin-Berlin**

09:00  Registration and Light Breakfast

10:00  Opening and KNIME Update
      Opening Remarks
      Michael Berthold (KNIME)
      What's New and Cooking
      Bernd Wiswedel and the KNIME Team

12:30  Lunch

13:30  Session 1
      H2O in KNIME: Integrating High Performance Computing in KNIME
      Jo-Fai Chow (H2O.ai) and Christian Dietz (KNIME)
      KNIME In Action: The Monster Model Factory
      Greg Lundrum (KNIME)
      Predictive Analytics of Hair Colorations
      Georg Knobbel (Herkel Beauty Care) and Iris Achterberg (Herkel Beauty Care)

15:00  Coffee Break

15:30  Session 2
      Blending Enterprise BI, Data Science, and Machine Learning
      Matthias Nolte (Berlin Sparkasse)
      Smart Data Collaboration Hub
      Rainer Dürr and Gerrit Burgerath (Landesbund der Deutschen Wirtschaft)
      KNIME and Teacher Bots: From Workflows to Apps
      Kathrin Meltcher, Vincenzo Tursi, Rosaria Silipo

17:15  End of Day 1

**Spring Summit Dinner**

17:45  Buses will shuttle between Hotel Berlin, Berlin-Berlin and Spindler & Klatt
      Hotel Berlin, Berlin-Berlin for return trip.

19:00  KNIME Spring Summit Dinner
      Spindler & Klatt – Köpenicker Str. 16, 10997 Berlin

21:00-24:00  Buses will shuttle between Spindler & Klatt and Hotel Berlin, Berlin-Berlin for return trip.

**Thursday March 8, 09:00 – 12:30, Rooms Berlin-Berlin 1 and Berlin-Berlin 4**

09:00  Registration and Light Breakfast

09:30  Parallel Track: Life Science Session 3a (Room Berlin-Berlin 1)
      Deep Learning: From Mastering the Game of GO to Revolutionizing Microscopy
      Florian Jüg (MPI-CBG Dresden)
      3D-e-Chem: Structural Cheminformatics & Discovery
      Chris de Graaf (VU Amsterdam)
      Guided Analytics in Action: KNIME Does Biomedical Research
      Jesny Prinz (KNIME)

12:30  Lunch

13:30  Session 3
      Parallel Track: Customer Intelligence 5
      KNIME In Action: GDPR, Taking a Provocative Stand
      Phil Winter (KNIME)
      Predicting Probability of Debt Recovery
      Elf Yigit Aycan and Mesut Böçü (Hayat Vasıf)
      Automating CRM, Marketing, and Analytic Edvinas Subaciaus (Zurigre Gaming)

15:00  Coffee Break

15:30  Session 4
      Integrated CADD Methods: A Cocktail of Strategies
      Louis Moretti (NucleicAcids)
      Exploring the Factory: Into the Gut of the Cell
      Greg Lundrum, Daria Goldmann, Anna March

17:15  End of Day 2

**Friday March 9, 09:00 – 14:00**

09:00  Registration and Light Breakfast

09:30-11:00  Workshops: Session 1

**Name of Workshop**

- Random Forest and Gradient Boosted Trees in KNIME
- Guided Analytics in Action
- KNIME Big Data Workshop
- de.NBI/CIBI Workshop: Analysis of Mass Spectrometry and Sequence Data with KNIME
- KNIME Deep Learning

**Instructors**

- Dean Abbott (SmarterHQ)
- Daria Goldmann and Jesny Prinz (KNIME)
- Tobias Köttler and Björn Lohrmann (KNIME)
- Tobias Köttler and Björn Lohrmann (KNIME)
- Julianus Pfeuffer (University of Tübingen)
- René Rahn (TU Berlin)
- Julian Uzorkew (University of Bochum)
- Alexander Fubruhn (University of Konstanz)
- Adrian Nemisch, Christian Dietz (KNIME)

**Room**

- Berlin-Berlin 1
- Berlin-London
- Berlin-Berlin 3
- Berlin-Peking
- Berlin-Montreal

11:00-11:30  Coffee Break

11:30-13:00  Workshops: Session 2

**Name of Workshop**

- Random Forest and Gradient Boosted Trees in KNIME
- KNIME Server Workshop
- KNIME Text Mining Workshop
- KNIME Image Processing Workshop

**Instructors**

- Dean Abbott (SmarterHQ)
- Jon Fuller and Iris Aδ (KNIME)
- Kilian Thiel (KNIME)
- Simon Schmid and Martin Horn (KNIME)

**Room**

- Berlin-Berlin 1
- Berlin-London
- Berlin-Berlin 3
- Berlin-Peking
- Berlin-Montreal

13:00-14:00  Lunch and End of KNIME Spring Summit 2018

© 2018 KNIME AG. All Rights Reserved.
Thank you and...

...have Fun!