Welcome & Opening

Michael Berthold

KNIME
Outline

• Why Summit?
• Recent KNIME Highlights
• A few Thoughts on KNIME and Data Science
• Summit Logistics
A Summit?
Once Upon a Time...

2006
KNIME v1.0
UGM 2007 (Konstanz): 12 Early Adopters and one Partner

UGM 2010, ..., 2014 (Zürich): >100 Fans, Users, Customers...

UGM 2015 & Summit 2016 (Berlin)
all of the above, and...
many curious data scientists
(≈250 attendees in 2016)

2006
KNIME v1.0
Spring Summit Attendees
125+ Fall Summit Attendees
A few Highlights of the Past Year
Growth in Konstanz
Growth in Berlin
Who is here?

From Zurich:
• Bernd Wiswedel
• Jon Fuller
• Greg Landrum

...Berlin:
• Tobias Kötter

...Konstanz:
• Heather Fyson
• Rosaria Silipo
• Peter Ohl
• Michael Berthold
• Christian Dietz

...and the US:
• Jay Diamond
• Arthur von Eschen
Who uses Analytics

Data Science?

Advanced Analytics
Analytics Personas

- New Algorithms come from scientists who write code:
  - R, Python, Java, C, ...
  - provides access as stand alone tool or as libraries/packages
- New Algorithms are used by data scientists:
  - knows Python, R, C, Java, SQL or ...
  - often struggles with data integration (nowadays: “data blending”)
- Analytical “Best Practices” are re-used by citizen data scientists:
  - maybe knows Java, Python, R, C, SQL or ...
  - struggles with backwards compatibility, reproducibility, ...
- But ultimately business analysts want to just use it!
  - don’t know (and should not know) C, R, Python, Java, or ...
  - want customized analytics... and keep bugging the data scientists.
Making Programming Scientists work with Data Scientists
Integrating R and Python

[Diagram showing integration of R and Python in KNIME flow]

© 2016 KNIME.com AG. All Rights Reserved.
Other Programming/Scripting Integrations

Java Integrations
- Java Snippet
- Java Snippet Row Filter
- Java Snippet Row Splitter

JavaScript Integrations
- JavaScript View
- JavaScript Table View
- JavaScript Scatter Plot
- JavaScript Lift Chart
- JavaScript Conditional Box Plot
- JavaScript ROC Curve
- JavaScript Line Plot
- JavaScript Box Plot
- JavaScript Bar Chart
- JavaScript Pie/Donut Chart

Misc Integrations
- External Tool
- External SSH Tool
- Generic Web Service Client
- SAS7BDAT Reader
...and the KNIME SDK
Data Scientists “program” visually
KNIME Analytics Platform Version 3.2
Visual Programming: Data Blending / SQL
SQL Modules

Database Connectors
- Database Connector
- PostreSQL Connector
- Hive Connector
- SQLite Connector
- Impala Connector
- Vertica Connector
- Node 22
- MySQL Connector

Database I/O
- Database Reader
- Database Connection Table Reader
- Database Looping
- Database Writer
- Database Update

Database Manipulation
- Database Table Connector
- Database Table Writer
- Impala Loader
- Database Table Selector
- Hive Loader
- Database Delete

Database Utility
- Apply-Binner
- Database Pivot
- Database Sorter
- Database Auto-Binner
- Database Column Filter
- Database GroupBy
- Database Numeric Binner
- Database Query
- Database Joiner
- Database Column Rename
- Database Row Filter
- Database Sampling
- SQL Extract
- SQL Inject

© 2016 KNIME.com AG. All Rights Reserved.
New Integrations: Big Data Data & Tool Blending
...and, of course: Partner Integrations!

Demo’ing at the Summit:

(Life Science integrations)

(Spotfire, SAP HANA integrations)

(Predictive Model Deployment)
Cloud Integrations

At KNIME®, we build software for fast, easy and intuitive access to advanced data science, helping organizations drive innovation. KNIME Cloud Analytics Platform brings the power and flexibility of the open source KNIME Analytics Platform to the cloud for the first time on Microsoft Azure.

With more than 1000 modules, hundreds of ready-to-run examples, community contributions, and tool integrations, KNIME Cloud Analytics Platform helps any data scientist solve their most complex data puzzles in a reliable and scalable framework with plenty of room to grow.

The combination of KNIME Cloud Analytics Platform on Microsoft Azure gives data scientists the power and flexibility to visually tackle complex data problems without the hassles of on-premise infrastructure or installation.

With KNIME Cloud Analytics With KNIME Cloud Analytics Platform, the only restriction is your creativity.
Citizen Data Scientists like Templates
Hiding Complexity (and code...)

• Not everyone should see all details
• Encapsulate “implementation”
⇒ KNIME (Wrapped) Metanodes
KNIME Metanodes: Controlling Parameters
KNIME Metanodes: Exposing select parameters
KNIME Metanodes: Encapsulating Data Science
Sharing Metanodes
Metanode Locking is part of Partner Productivity Tools. (Requires Commercial License)
Business Analysts just want to use it.
Classic Business Analysts

Business Analysts currently have access to:

• BIRT Report Engine Integration
• Integrations with Visualization Partners (embedded analytics)

• KNIME WebPortal for supervised analytics
KNIME Web Portal: Deploying Configurable Workflows
Is that “Analytics for the Masses”?
Analytics for Everybody

Requires

• pre-built analytics workflows...
• ...flexible enough to incorporate user context

→ Analytics Applications require Guidance!
Guided Analytics

Analytics Consumer

KNIME WebPortal

Data Scientist

Explanations

Analytics with Interaction Points

Start here

Feedback

© 2016 KNIME.com AG. All Rights Reserved.
Guided Analytics

Later today:

Rosaria Silipo and Greg Landrum: “Data Science for Everybody”
The Fall Summit
Thursday, September 15

9:00 am  |  Registration and Breakfast: Atrium

---

**Opening and KNIME Session**

10:00 am  |  Welcome and Introduction
           |  Michael Berthold (KNIME)

10:30 am  |  What’s New in KNIME
           |  Bernd Wiswedel (KNIME)

11:30 am  |  What’s Cooking in the KNIME and Big Data Labs
           |  Bernd Wiswedel & Tobias Kötter (KNIME)

12:30 pm  |  Lunch Break: Atrium

---

**Session I**

1:30 pm   |  Open Source Innovation in a Closed Source World
           |  Andrew Ruegger & Greg Smith (GroupM/Catalyst)

2:00 pm   |  The Universal Model: Realizing Business Value through Near Real-Time Decision Support Systems
           |  Ramya Narayanaswamy & May Tan (Comcast)

2:30 pm   |  Afternoon Break: Atrium
### Thursday, September 15

#### Session II

<table>
<thead>
<tr>
<th>Time</th>
<th>Title</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>3:00 pm</td>
<td>Extracting Truths from Social Media</td>
<td>Benjamin Spiegel (MMI Agency)</td>
</tr>
<tr>
<td>3:30 pm</td>
<td>Weaponizing Sales and Marketing with Predictive Real-Time Lead Ranking</td>
<td>Jere Helenius (Palo Alto Networks)</td>
</tr>
<tr>
<td>4:00 pm</td>
<td>What's Love Got To Do With It? Analyzing Tweets about Online Dating</td>
<td>Cathy Pearl</td>
</tr>
<tr>
<td>4:30 pm</td>
<td>Short Break</td>
<td></td>
</tr>
</tbody>
</table>

#### Special Session

<table>
<thead>
<tr>
<th>Time</th>
<th>Title</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>4:45 pm</td>
<td>Data Science for Everybody</td>
<td>Rosaria Silipo &amp; Greg Landrum (KNIME)</td>
</tr>
</tbody>
</table>

#### Flying Buffet

<table>
<thead>
<tr>
<th>Time</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>5:30 pm</td>
<td>Networking: Atrium</td>
</tr>
</tbody>
</table>
Friday, September 16

9:00 am  Registration and Breakfast: Atrium

Session III

9:30 am  From Corkscrew to Swiss Army Knife: The evolving role of KNIME at DNS
          Brock Luty (Dart NeuroScience)

10:00 am  KNIME as a Platform for Integrated Hit Finding at Novartis
           Patrick Lee (Novartis)

10:30 am  Predictive Modeling of On-And-Off Target Bioactivities - A Data Vendor Perspective
           Tim Hoctor (Elsevier)

11:00 am  Morning Break: Atrium

Session IV

11:30 am  Open Source Tool Blending: Image Analysis in KNIME
           Curtis Rueden (University of Wisconsin-Madison) & Christian Dietz (University of Konstanz)

12:00 pm  How Big Data Analytics Powered by KNIME Can Transform a 120 Year Non-Profit NGO
           Nathaniel Lin (National Fire Protection Association)

12:30 pm  Lunch Break: Atrium
Friday, September 16

Session V

1:30 pm  Using KNIME agile prototyping to establish PoV to project sponsors
Sameer Bhatnagar (Consultant)

2:00 pm  When the Old meets the New: How Advanced Data Analytics is Transforming Manufacturing
Rumi Ghosh (Bosch)

2:30 pm  Transforming Under Armour's Market Execution through Geospatial Analytics
Ryan Keane (Under Armour)

3:00 pm  Afternoon Break: Atrium

Special Lecture

3:30 pm  Measuring Variable Importance with Target Shuffling
Dean Abbott (Abbott Analytics)

End of Fall Summit

4:30 pm  Get-together: Spark Social SF
...a little peak into the future...

coming to a workflow near you on December, 6, 2016.
Thank You,
have Fun,...
and see you in Berlin!

Mark your Calendars (and book flights):

**KNIME Spring Summit**
Berlin, Germany
March 15-17, 2017
The KNIME® trademark and logo and OPEN FOR INNOVATION® trademark are used by KNIME.com AG under license from KNIME GmbH, and are registered in the United States.