What’s new...

Bernd Wiswedel
KNIME.com AG, Zurich, Switzerland
Two feature releases last year: 2.8 & 2.9

Documented in Changelog, “What’s new summary” and as video on YouTube
What's new page on knime.org

What's New in KNIME 2.9

Highlights
- Rule Based Framework Extended and Improved
- Column Selection Framework Extended and Improved
- Better Model Training and Evaluation
- Werkbank Monitoring, Performance, and Administration
- Date Generation
- Other New Features and Nodes
- Completely changed log file

Rule Based Framework Extended and Improved

New Rule Engine Node

New, more powerful, more flexible, and more user-friendly Rule Engine Node.

The configuration window of the new Rule Engine node follows the same GUI as the string manipulation node and as the math formula node. This new user interface allows for an easier inclusion of flow variables and data columns in the rule editing frame and for replacing existing columns in the output data table.

PMML Ruleset Editor

New Node

This new node is based on the framework of the new Rule Engine node (see above).

It works like the new Rule Engine node, producing the resulting values as well as the rule set as a PMML model at the output ports.
Learning Hub & Youtube
**Changelogs**

### Changelog v2.8.0

**Changes from v2.7.4 to v2.8.0**

**Release Date:** July 24, 2013

**New Nodes:**
- New Node 4137: Open Street Map Plugin
- New Node 3969: Column Append (date joins)
- New Node 4208: Lag Column - Append new column after lagging
- New Node 4000: Render to Image (pdf export)
- New Node 3886: Multi-Sheet XLSS Writer Output
- New Node 2928: Read XLSS sheet names
- New Node 3943: URI Port to Variable
- New Node 4081: PMML Support for the Encyclopedia
- New Node 4003: Many2One with PMML support
- New Node 4169: Column Filter node with PMML support
- New Node 4134: R Source
- New Node 3065: Bootstrap sampling
- New Node 4265: Delete Files
- New Node 4268: Create Directory

### Changelog v2.9.0

**Changes from v2.8.2 to v2.9.0 (see highlight summary)**

**Release Date:** December 05, 2013

**New Nodes:**
- New Node 3867: Send Mail node with HTML option and file attachments
- New Node 4741: Save Workflow on execution
- New Node 4269: Row Filter based on Rule Engine functionality
- New Node 4270: Row Splitter based on Rule Engine functionality
- New Node 4531: Cross Joiner
- New Node 4271: Rule Engine node for variables
- New Node 4414: RuleSet Editor and RuleSet Predictor with PMML ports
- New Node 4293: Active Branch Inverter node to toggle active status of a branch
- New Node 4345: Data Generation: Counter Generation adds a simple counter to a table
- New Node 4559: Data Generation: Time Series Generator node
- New Node 4871: Network Mining: Node that converts a distance matrix to a network
- New Node 4704: KNIME Labs: PMML Classifier, Cluster Assigner, and Regression Predictors
- New Node 3818: SDF Stripper node
- New Node 4064: Extract Content Properties
Changelogs

KNIME Desktop 2.8/2.9
• New Nodes: 43
• New Extensions: 2
• New Features: 122

KNIME Server/Teamspace 3.7/3.8
• New Features: 33
Outline

New features demos
Presented by the developers
Selected workflows later available via publicserver.knime.org
Outline I - Changes in the Desktop

- R Statistics Integration
- Open Street Map Plugin
- “Try-Catch” nodes, error handling
- Improved Excel Reader/Writer
- New Nodes: Rule Editor and Rule Filter
- Loops: Recursive Loops, Feedback Loops
- Text Mining: New Nodes & major speedup
- Other Changes
Outline II - Changes in the Server

• Versioning: Archive workflows, rollback
• Usability & Robustness
• Webportal: Custom URLs & special triggers
Outline I - Changes in the Desktop

- Heiko Hofer: R Statistics Integration
- Open Street Map Plugin
- “Try-Catch” nodes, error handling
- Improved Excel Reader/Writer
- New Nodes: Rule Editor and Rule Filter
- Loops: Recursive Loops, Feedback Loops
- Text Mining: New Nodes & major speedup
- Other Changes
Outline I - Changes in the Desktop

- R Statistics Integration
- Martin Horn: Open Street Map Plugin
- “Try-Catch” nodes, error handling
- Improved Excel Reader/Writer
- New Nodes: Rule Editor and Rule Filter
- Loops: Recursive Loops, Feedback Loops
- Text Mining: New Nodes & major speedup
- Other Changes
Open Street Map
Open Street Map
Open Street Map
Open Street Map
Open Street Map
Open Street Map
Open Street Map
Open Street Map
Open Street Map

http://www.openstreetmap.org
http://josm.openstreetmap.de
Open Street Map
Outline I - Changes in the Desktop

- R Statistics Integration
- Open Street Map Plugin
- “Try-Catch” nodes, error handling
- Improved Excel Reader/Writer
- New Nodes: Rule Editor and Rule Filter
- Loops: Recursive Loops, Feedback Loops
- Text Mining: New Nodes & major speedup
- Other Changes
Outline I - Changes in the Desktop

• R Statistics Integration
• Open Street Map Plugin
• “Try-Catch” nodes, error handling
• Thomas Gabriel: Improved Excel Reader/Writer
• New Nodes: Rule Editor and Rule Filter
• Loops: Recursive Loops, Feedback Loops
• Text Mining: New Nodes & major speedup
• Other Changes
Outline I - Changes in the Desktop

- R Statistics Integration
- Open Street Map Plugin
- “Try-Catch” nodes, error handling
- Improved Excel Reader/Writer
- Thorsten Meinl: New Nodes: Rule Editor and Rule Filter
- Loops: Recursive Loops, Feedback Loops
- Text Mining: New Nodes & major speedup
- Other Changes
Rule Engine

- User-defined IF-THEN rules
- similar to SQL WHERE statements
- Improved functionality and dialog in 2.9
- New Rule-based Row Filter & Splitter nodes
- New PMML Ruleset Editor & Predictor nodes
Rule-based Row Splitter

- Uses same rules, but with TRUE/FALSE outcomes
- First matching rule determines output
  - No match is FALSE outcome

Expression

1. $\text{Distance} < 3 \Rightarrow \text{FALSE}
2. $\text{Distance} > 10 \text{ AND } \text{name length} < 30 \Rightarrow \text{FALSE}
3. TRUE $\Rightarrow$ TRUE

TRUE to first, FALSE to second output table
PMML Ruleset Editor

- Rules compatible with PMML Rulesets
- No string matching
- No flow variables
- Only constants in outcomes
- Outputs classification and PMML model
Math Formula

- Same dialog concept
- Support for Flow variables
Outline I - Changes in the Desktop

- R Statistics Integration
- Open Street Map Plugin
- “Try-Catch” nodes, error handling
- Improved Excel Reader/Writer
- New Nodes: Rule Editor and Rule Filter

**Iris Adä: Recursive Loops, Feedback Loops**

- Text Mining: New Nodes & major speedup
- Other Changes
Rule Engine

- User-defined IF-THEN rules
- similar to SQL WHERE statements
- Improved functionality and dialog in 2.9
- New Rule-based Row Filter & Splitter nodes
- New PMML Ruleset Editor & Predictor nodes
Rule-based Row Splitter

- Uses same rules, but with TRUE/FALSE outcomes
- First matching rule determines output
  - No match is FALSE outcome
PMML Ruleset Editor

• Rules compatible with PMML Rulesets
  • No string matching
  • No flow variables
  • Only constants in outcomes

• Outputs classification and PMML model
Math Formula

- Same dialog concept
- Support for Flow variables
Recursive Looping

Iris Adä
Recursive Loops
1. Iteration
1. Iteration

Collecting
Collects data for the outport

Recursive Loop Start

Recursive Loop End

Data Handling

Recursive
Collects data for the loop start
2. Iteration
n. Iteration
2 Ports
Recursive Looping Demo
Outline I - Changes in the Desktop

- R Statistics Integration
- Open Street Map Plugin
- “Try-Catch” nodes, error handling
- Improved Excel Reader/Writer
- New Nodes: Rule Editor and Rule Filter
- Loops: Recursive Loops, Feedback Loops
- Kilian Thiel: Text Mining - New Nodes & major speedup
- Other Changes
Text Processing Enhancements
File Store Cells

New cells storing documents

- Preferences->KNIME->Textprocessing->Storage
- Faster than Blob Cells
- New de/serialization
Direct Preprocessing

• Apply preprocessing nodes directly on list of documents
• Create bag of words afterwards
Direct Preprocessing
Outline I - Changes in the Desktop

- R Statistics Integration
- Open Street Map Plugin
- “Try-Catch” nodes, error handling
- Improved Excel Reader/Writer
- New Nodes: Rule Editor and Rule Filter
- Loops: Recursive Loops, Feedback Loops
- Text Mining: New Nodes & major speedup
- Other Changes
Unifying learner/predictor outputs - 1

Model Parameters easily accessible

Standardized predictor output
Learner Nodes to output parameters, eg.

- Cluster centers in k-Means
- Regression Coefficients

Unifying learner/predictor outputs - II
Workflow Control

Save As...

Ability to execute “red” nodes
KNIME Explorer

KNIME Explorer with data files, meta nodes, ...

KNIME Protocol to read workflow relative paths:

knime://knime.workflow/path.csv
Column Selections - More control

Various nodes use multi column selection:
  Column Filter, Round Double, Collection Create, Splitter, ...

Column Selection Panel with filtering based on name patterns (wildcards) or column types
New Nodes

Too many to mention all ... but...

- Lag Column: A must in time series analysis
- Constant Value Column
- Send Email
- Bootstrap Sampling
- Cross Joiner
Outline I - Changes in the Desktop

- R Statistics Integration
- Open Street Map Plugin
- “Try-Catch” nodes, error handling
- Improved Excel Reader/Writer
- New Nodes: Rule Editor and Rule Filter
- Loops: Recursive Loops, Feedback Loops
- Text Mining: New Nodes & major speedup
- Other Changes
Outline II - Changes in the Server

- Versioning: Archive workflows, rollback
- Usability & Robustness
- Webportal: Custom URLs & special triggers
Outline II - Changes in the Server

Thorsten Meinl:

• Versioning: Archive workflows, rollback
• Usability & Robustness
• Webportal: Custom URLs & special triggers
KNIME Enterprise Server Features in 3.7 & 3.8
New repository structure

- All information is now stored on disk
- Backup & restore much easier
- Improved performance with large repositories
  - ~4000 workflows scanned in 2.5 seconds
  - Client shows repository contents almost immediately
- Auto-conversion of old repositories
Versioning overview

• Workflows and data files can be versioned
• A new version or snapshot can be created before overwriting an item on the server
• Older version can be restored (replacing the current visible version), or...
• Downloaded into the local workspace
Recycle bin overview

• Deleted items are now moved into a recycle bin by default
• Contents of recycle bin are visible in the new Server Recycle bin view
• Deleted items can be restored on the server, or...
• Downloaded into the local workspace
Permission inheritance

- Any item can now inherit permissions from its parent workflow group.
Permission inheritance II

- Inheritance is dynamic...
  - Permission changes on the parent are immediately reflected in the child
- ... and transitive
- Default permissions for new items are "inherit from parent"
Mountpoint editor I

- Improved dialog for all mountpoint types
Mountpoint editor II

- Default mount id provided by server
HTML reports I

- Reports can now be generated in HTML format
- Images are included in the HTML enabling previews in notification emails
Subset Details for: Iris setosa

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sepal.Length</td>
<td>5.01</td>
<td>0.35</td>
<td>4.3</td>
<td>5.8</td>
</tr>
<tr>
<td>Sepal.Width</td>
<td>3.43</td>
<td>0.38</td>
<td>2.3</td>
<td>4.4</td>
</tr>
<tr>
<td>Petal.Length</td>
<td>1.46</td>
<td>0.17</td>
<td>1</td>
<td>1.9</td>
</tr>
<tr>
<td>Petal.Width</td>
<td>0.25</td>
<td>0.11</td>
<td>0.1</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Sepal Dimensions

Petal Dimensions

Created with KNIME Report Designer. Provided by KNIME.com AG, Zurich, Switzerland
RMI improvements

• Inactive jobs are...
  - Swapped to disk after 60 minutes
  - Discarded after one week

• RMI instance can be restarted without server restart
  - E.g. to activate new nodes
Outlook

- Improved versioning
  - Associate jobs with workflow version
  - Unique, increasing id for versions
- Improved metanode linking
- GUI improvements in KNIME Explorer
  - Background operations
  - Open server workflow with double-click
Outline II - Changes in the Server

- Versioning: Archive workflows, rollback
- Usability & Robustness
- Christian Albrecht: Webportal: Custom URLs & special triggers
URL Handling

- Path to workflows and jobs
- Parameters:
  - Single-Mode (&single)
  - Auto-run (&run)
  - Show sub-workflow tree (&root=path/to/root)
  - Setting Workflow Variables (&wv:name=value)
  - Setting Credentials (&cr:name=user:password)
  - Notification Settings (&emails=sample@knime.com)
  - Report Formats (&formats=pdf,html)
Additional custom settings

- Additional title label
- Disable HTML report preview
- Disable warning messages
Summary: What’s new

Discussed KNIME Usage Examples
→ check “Examples” Server for even more

New functionality constantly added, thanks to community, partners and customers

And more is coming...