



# The 4<sup>th</sup> KNIME User Group Meeting - Welcome & Introduction -

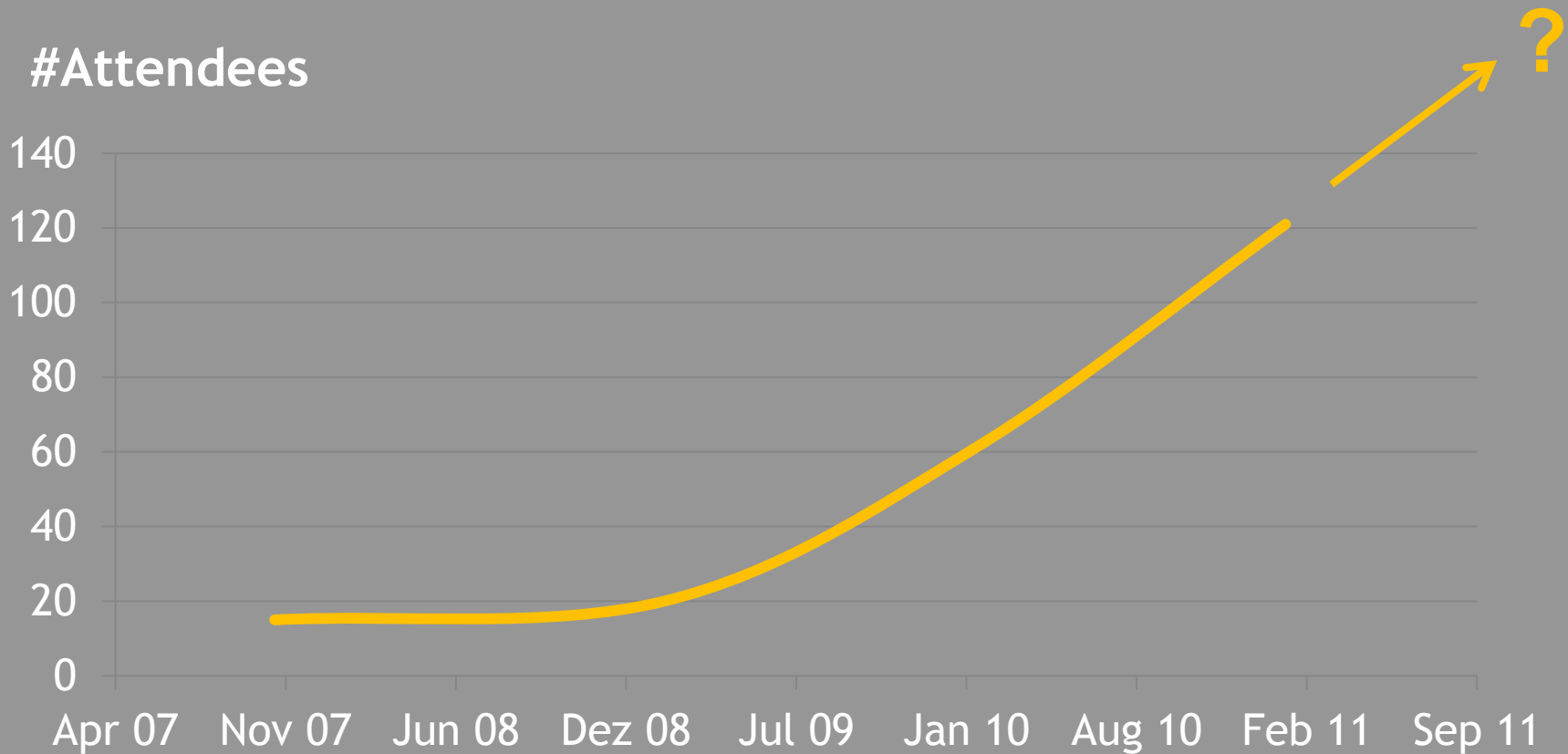
*Michael R. Berthold*  
*KNIME.com GmbH, Zurich, Switzerland*



## KNIME Meetings

- 1<sup>st</sup> UGM (Konstanz): November 2007
- 2<sup>nd</sup> UGM (San Francisco): February 2009
- 3<sup>rd</sup> UGM (Zürich): February 2010
  - Life Science Day (Boston): August 2010
- 4<sup>th</sup> UGM (Zürich): February 2011
  - Life Science Day (SF): Summer 2011

# UGM Attendance





## 2010 Highlights

# 2010 Highlights: The New Website



## KNIME Enterprise Server

Enterprise Server provides additional capabilities to use KNIME in a professional corporate environment.

[/ More information.](#)



+++++ **Only a few seats left!** +++++

Next week the doors will be opened for the **4th KNIME User Group Meeting in Zurich, Switzerland** from February 28 to March 4, 2011. Hurry up - more than 100 KNIME users have already registered! [Read more...](#)

## KNIME - Professional Open-Source Software

KNIME (Konstanz Information Miner) is a user-friendly and comprehensive open-source data integration, processing, analysis, and exploration platform. From day one, KNIME has been developed using rigorous software engineering practices and is currently being used actively by over 6,000 professionals all over the world, in both industry and academia.

### Forum & Documentation



### News

[Dymatrix introduces KNIME for DynaMine Data Mining Automation and DynaCampaign...](#) 2011/02/17

[KNIME Newsletter Vol. 1, No. 1 - January 2011](#) 2011/01/31

[KNIME Version 2.3.1 released](#) 2011/01/21

[4th KNIME Users Group Meeting from February 28 to March 4 in Zurich, Switzerland](#) 2011/01/20

[KNIME Version 2.3.0: One more thing...](#) 2010/12/24 [more](#)

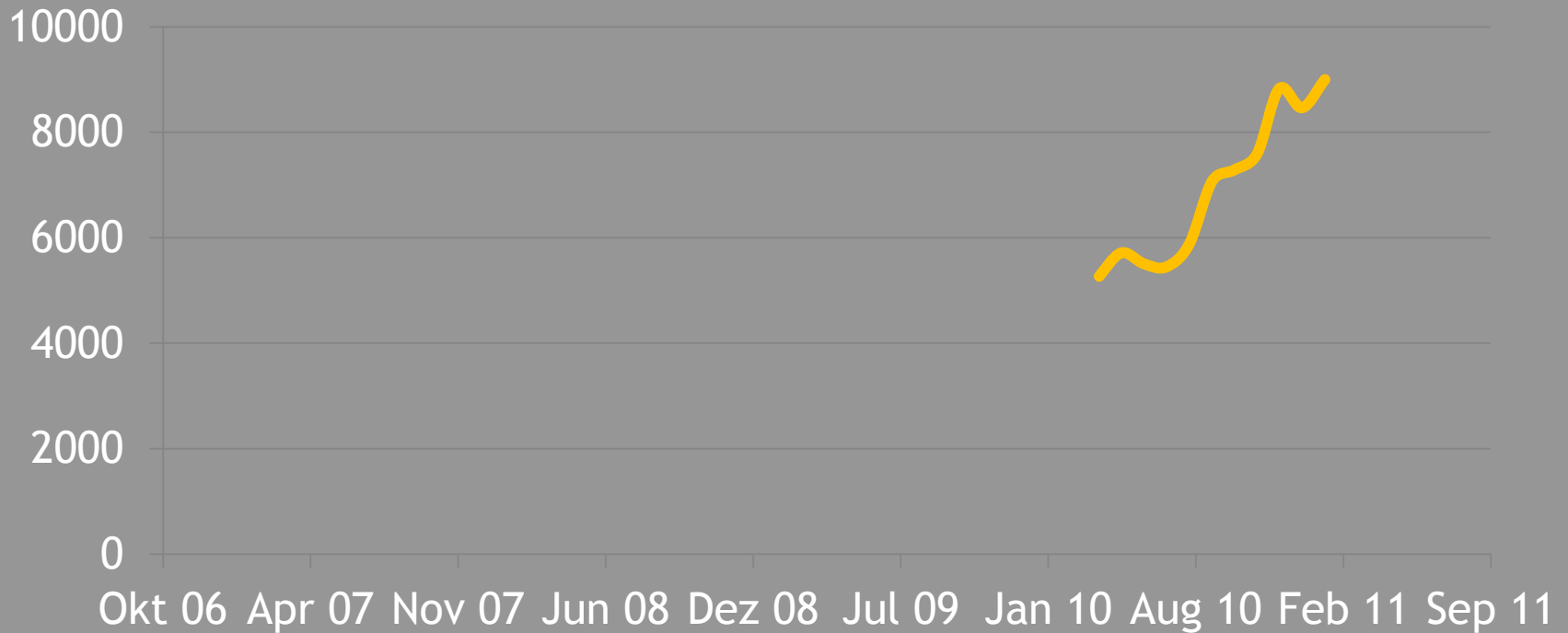
### Upcoming events

[4th KNIME Users Group Meeting and Workshop](#)



# (Unique) KNIME Website Visitors

Visitors / Month





## 2010 Highlights: Enterprise Products

- KNIME (Report) Server
  - Share and Deploy Workflows and Reports
- KNIME Cluster Execution
  - Run Workflows on your Cluster
- Already running at numerous customer sites



## 2010 Highlights: Community Site

- Central Update Site for Community Nodes
  - Scripting Support
  - Image Processing
  - More cheminformatics functionality

# 2010 Highlights: KNIME News

## Quarterly Newsletter

### KNIME Newsletter

January 2011 - Vol. 1, No. 1

**Welcome to the first KNIME Newsletter!**

The KNIME community has grown quickly over the last few years. From a few hundred downloads of KNIME 1.0 in July 2008 there are now thousands of people around the world actively using KNIME. The KNIME Newsletter is our newest vehicle to support the growing community—every quarter we will report on a particularly interesting application of KNIME in the real world and complement this with a series of tips and tricks around the practical usage of KNIME.

We hope you enjoy reading this and future versions of KNIME News!

The KNIME Team

**4th KNIME User Meeting**

**Dear Group Meeting 2010!**

Meet the KNIME Users and the entire KNIME Team!

**March 2-4, 2011**  
 Technopark  
 Zurich, Switzerland  
[www.knime.org/um](http://www.knime.org/um)

If you are interested to learn more about how to use KNIME and KNIME Reporting you can attend for the associated training courses.

**KNIME User Training**  
**Feb. 28 - March 1, 2011**  
 Technopark  
 Zurich, Switzerland

**Reports with KNIME, a success story**

People working in the corporate world and, in particular, people from the financial world, will know just how many reports are required every year, every month, and even every day for the company to work properly. Construction across departments and between employees and management is crucial if the bank's activities are to run smoothly. Some reports are required on demand while others are needed on a regular basis depending on the specific nature of their content. An automatic scheduling solution is often required for regularly called reports in order to free up some of the company resources. Reports must also be repeatable. Business people in banks and other companies need reports that integrate data from diverse data sources. This frequently poses a technical problem. Data is often spread across one or more databases, it may exist aheads of different structures and across flat files, and maybe even employee written notes.

Bank Frey is a medium-sized Swiss private bank. It was founded in 2000 with the strong commitment to secure and grow wealth for generations to come. Thanks to its talented team of top professionals with an entrepreneurial focus, Bank Frey has grown rapidly in the last 10 years. This rapid growth has made evident the need for a reliable and scalable report platform to cover all internal and external communication requirements.

Bank Frey was on the lookout for a Business Intelligence (BI) and reporting tool with low license costs and minimum training requirements that would work across a number of factors. Greater BI integration was also a goal. Bank Frey needed the scalability to Swiss federal banks' growing amounts of data. After a rapid evaluation of the BI tools available on the market, Bank Frey decided to test the KNIME Reporting tool.

Unlike many other BI tools, KNIME, being open source, can be fully tested free of charge. As a tool was to be selected, KNIME it and one of its practical interfaces is intuitive and does not require any previous knowledge of script or language programming. The full range of BI functionalities is available, spanning from the calculation of classic statistical variables to data manipulation functions to integration and re-arrangement. Some basic knowledge of statistics is still needed, but this provided the transition to KNIME is quick and painless.

After only a few months of work, Bank Frey has set up a simple Data Warehousing (DW) solution, including a number of reports run on either a regular monthly or daily basis and further reports that are run on demand. The DW workflow integrates data from Excel files, text files and databases and stores them in a centralized location. The Report workflow then performs the required data selection and writes the results into a report with a professional graphic layout. The DW and the

**Bank Frey**



Example of the reports running at Bank Frey.

**KNIME 2.3 is out!**

Dear KNIME Users, we are excited to announce the new KNIME 2.3. You can download it from [www.knime.org/download](http://www.knime.org/download) and start using your data today!



## Announcements

### Welcome to the first KNIME Newsletter!

The KNIME team has been working for a few years. From a few hundred downloads to a growing community—every quarter we will report on our progress and complement this with a series of tips and tricks.

## User Reports

We hope you are enjoying this and future versions of KNIME News!

The KNIME Team

### Reports with KNIME: a success story

People working in the corporate world and, in particular, people from the financial world, well know just how many reports are required every year, every month, and even every day for the company to work properly. Communication across departments and between employees and management is crucial if the bank's activities are to run flawlessly.

**"Data from different sources is consolidated and is now available at any moment for reporting and analytics."**

Some reports are required

order to free up some of the company resources. Reports must also be repeatable. Business people in banks and other companies need reports that integrate data from diverse data sources. This frequently poses a technical problem. Data is often spread across one or more databases, a few Excel sheets of different structures here, some text files there, and maybe even employees' written notes.

Bank Frey is a medium-sized Swiss private bank. It was founded in 2000 with the strong commitment to secure and grow wealth for generations to come. Thanks to its talented team of top professionals with an entrepreneurial focus, Bank Frey has

Bank Frey was on the lookout for a Business Intelligence (BI) and reporting tool with low license costs and minimum training requirements that would work across a number of heterogeneous data sources and provide the scalability to handle growing amounts of data.

After a rapid evaluation of the BI tools available on the market, Bank Frey decided to test the KNIME Reporting tool.

Unlike many other BI tools, KNIME, being open source, can be fully tested free of charge. All it



**Gregor Bienz**  
CEO  
Bank Frey  
Zurich  
Switzerland

KNIME news

### 4th KNIME User Meeting

#### User Group Meeting (UGM)

Meet other KNIME Users and the entire KNIME Team!

**March 2—3, 2011**  
Technopark  
Zurich, Switzerland

[www.knime.org/ugm](http://www.knime.org/ugm)

If you also want to learn more about how to use KNIME and KNIME Reporting, you can enroll for the associated training courses.

#### KNIME User Training

**Feb. 28—March 1, 2011**  
Technopark  
Zurich, Switzerland

### Next Developer Training

During this course you will be

## What's new in KNIME 2.3

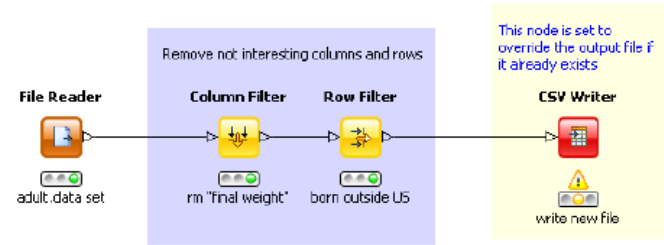
On the evening of December 5, children in northern Europe put their shoes outside the front door. They know that overnight, Nikolaus (Santa Claus) will stop at every door to fill the shoes of children who have been good with all sorts of sweets for them to find in the morning.

Late last year Nikolaus has taken a wide detour and stopped at the doors of all KNIME users to fill their shoes with a new more powerful version of KNIME: KNIME 2.3. We have largely retained the familiar look and feel of KNIME in this release and added a few new interesting features that simplify and empower our daily data analysis work. This section of the newsletter is dedicated to the new features of version 2.3.

KNIME 2.3 can be downloaded from

[www.knime.org/download](http://www.knime.org/download)

Shift+F7	executes all executable nodes
F7	executes selected and executable nodes
Shift+F9	Cancels all running nodes
F9	Cancels selected and running nodes
Ctrl+S	saves the workflow



This workflow is my first KNIME workflow. It reads data, removes uninteresting columns and rows from the data table and writes the result to a CSV file.

# Tips & Tricks

Use workflow annotations to comment on a node or group of nodes. Just right-click on the workflow editor and select "New Workflow Annotation". Add your explanatory comment, resize the annotation box to be clear which group of nodes it refers to, and there you have the explanation of this part of your workflow! The result is an improved overview of the whole workflow task.

**Hotkeys**  
For keyboard lovers and in general for all those users suffering from carpal tunnel syndrome, most KNIME commands can now run via hotkeys. All hotkeys are listed in the KNIME menus or in the tooltip messages at the side of their corresponding commands.

## Tips and Tricks

### Label Data Sets

Very often we need to concatenate a number of data sets to feed the data analysis cycle. However, once the data analysis has been performed, we might want to recreate the original data subsets exactly as they were.

Recently, for example, I created a data set with the game scores from the 2006 Football World Cup and a data set with the game scores of the 2002 Football World Cup from the data available on the FIFA web site ([www.fifa.com](http://www.fifa.com)). In order to perform an analysis on both world cups' scores, I concatenated the two data sets into one big data set. I also wanted to display the scores of each world cup, 2002 and 2006, in two separate report tables.

Before embarking on a cumbersome operation to identify the original data set of each data row by parsing the game date, I opted for a simpler solution: I assigned the world cup's year as an additional label to the data rows of each data set before concatenation.

To assign labels to data rows I often use the "default" label of the "Rule Engine" node. So, I used two "Rule Engine" nodes to create a new column "year" with default label "2002" in one data set and default label "2006" in the other data set. After



**Rosaria Silipo**  
Data Mining Consultant  
bizzeps AG  
Zurich, Switzerland

## 2010 Highlights: KNIME Press

- First KNIME Guide:  
“KNIME Beginner’s  
Luck”  
by Dr. Rosaria Silipo  
(Dec 2010)
- eBook with possible  
2 year subscription





## 2011 Preview: KNIME Press

- ...watch out for the SAS-to-KNIME guide...

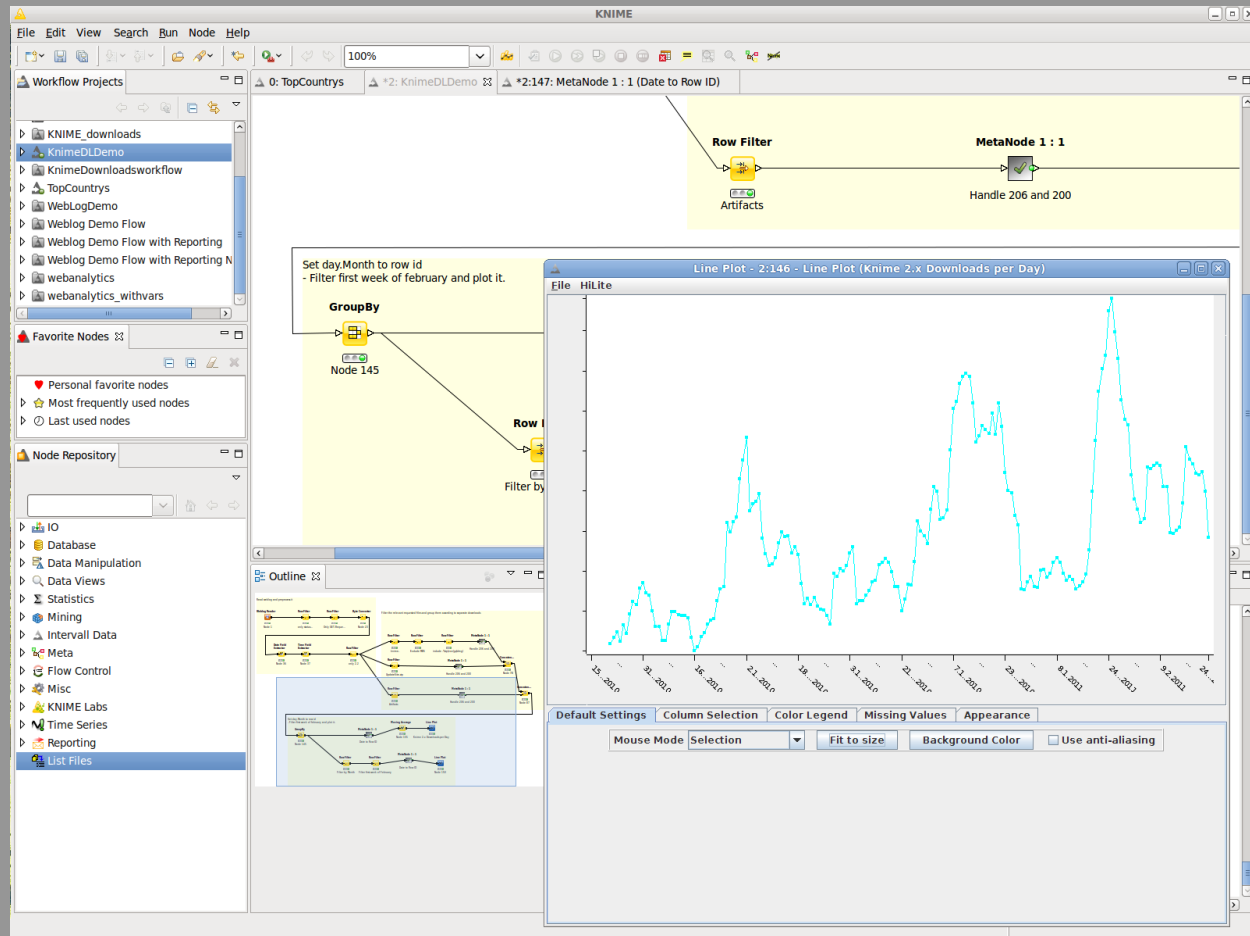


## 2010 Highlights: KNIME 2.3

and, of course, continued development of  
KNIME itself!

..."KNIME Today..." talk right after this...

# KNIME Downloads

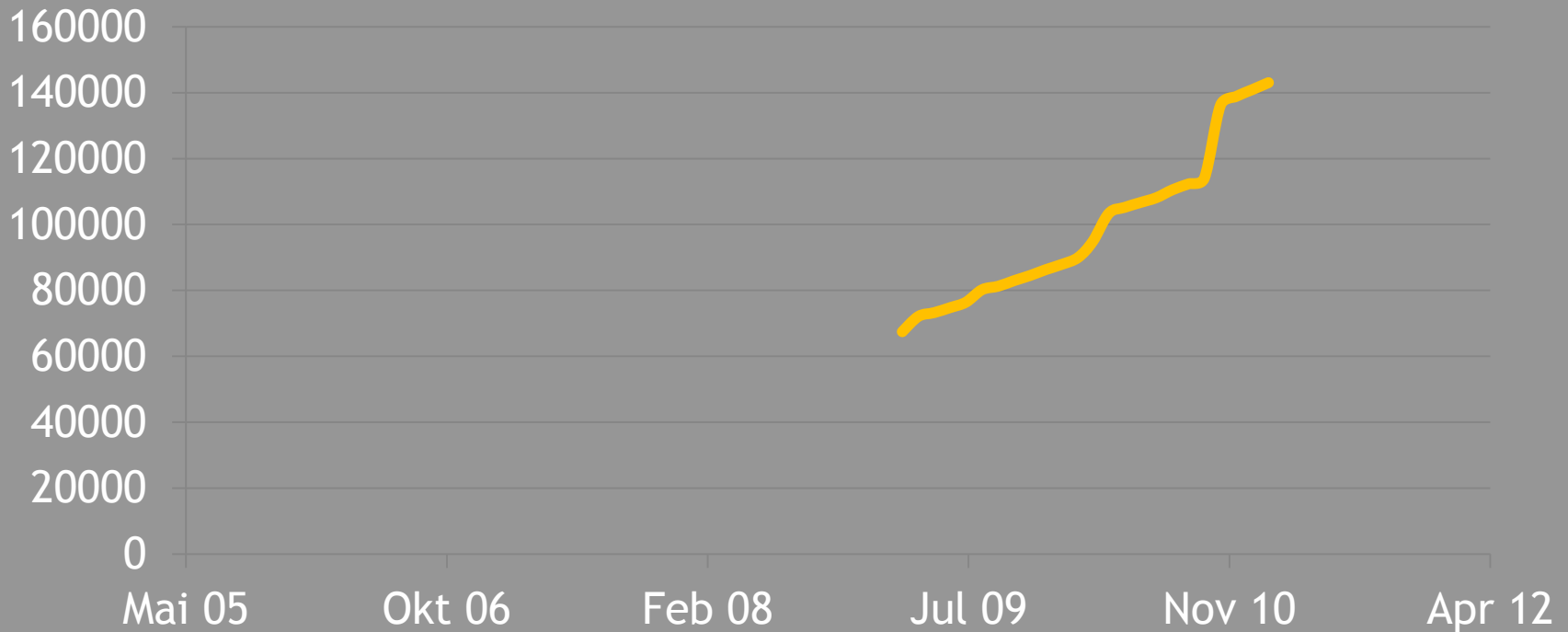


The screenshot displays the KNIME software interface. The main workspace shows a workflow with the following nodes: **Row Filter** (Artifacts), **MetaNode 1 : 1** (Handle 206 and 200), **GroupBy** (Node 145), and **Filter By** (Row ID). A text box above the GroupBy node contains the instructions: "Set day.Month to row id - Filter first week of february and plot it." The **Outline** panel at the bottom left shows a hierarchical view of the workflow nodes. A **Line Plot** window is open, titled "Line Plot - 2:146 - Line Plot (Knime 2.x Downloads per Day)". The plot shows a time series of data points connected by a red line, with the x-axis representing dates from 01-01-2017 to 24-02-2017. The plot shows a clear upward trend in downloads over time, with a significant peak in late February. The plot window includes a toolbar with options like "Default Settings", "Column Selection", "Color Legend", "Missing Values", and "Appearance".



# Overall KNIME Downloads to Date

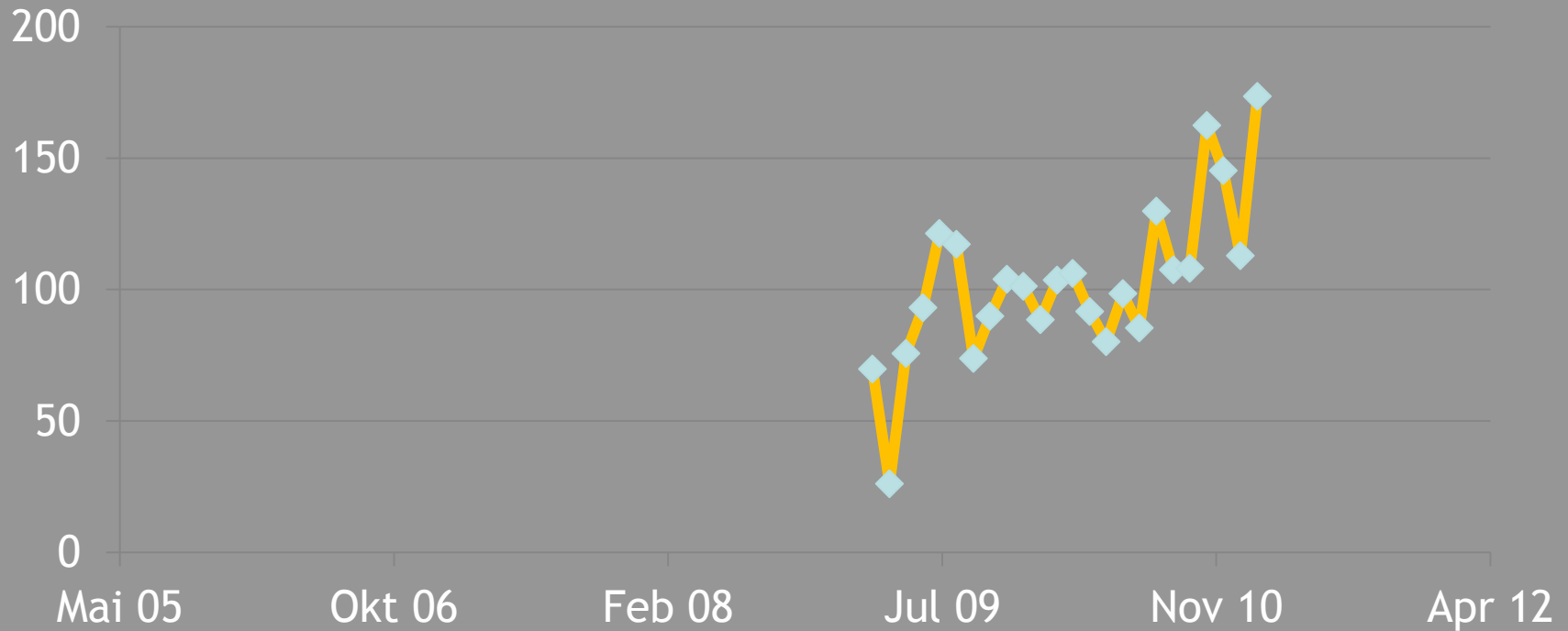
total Downloads





# Average KNIME Downloads per Day

Daily Downloads



## Miscellaneous 2010 Highlights

- 2010 Gardner Cool Vendor (“Business Intelligence and Analytics”)

A blue and white badge with rounded corners. The top half is white with the text 'Gartner | 2010' in blue. The bottom half is blue with the text 'COOL VENDOR' in white.

**Gartner | 2010**  
**COOL VENDOR**

- (Integration) Partner in EU Colipa Project “Toxicology in Cosmetics” (COSMOS)



# KNIME UGM Agenda

## **KNIME Users Meeting Day 1 – General Sessions**

Wednesday March 2, 10:00 am – 5:00 pm, Technopark Zurich, Auditorium

9:00 am            **Registration and Breakfast**

### **Session 1: KNIME Session**

10:00 am        **Welcome and Introduction**  
*Michael Berthold, KNIME*

10:30 am        **KNIME Today: New Features of Version 2.3**  
*Bernd Wiswedel, KNIME*

11:15 am        **Coffee Break**

11:30 am        **KNIME Labs – See What's Cooking**  
*Thomas Gabriel, KNIME*

12:30 pm        **Lunch Break**

# KNIME UGM Agenda

## Session 2: KNIME Applications

- 1:30 pm **Keynote Real-World Advanced Customer Intelligence**  
*Phil Winters, Peppers & Rogers Group*
- 2:15 pm **The use of KNIME in the development of IR-based indices of soil condition across landscapes**  
*Tor Vagen, World Agroforestry Centre*
- 2:45 pm **Explorative Analysis and Visualization of Multi-Dimensional and Geo Related Fuel Survey Data with KNIME**  
*Alexander Warta, Robert Bosch GmbH*
- 3:15 pm **Survey Data Collection**  
*Rosaria Silipo, bizzeps AG*
- 3:30 pm **Coffee break**
- 4:00 pm **Pervasive DataRush: Parallel data analysis with KNIME**  
*Jim Falgout, Pervasive Software*
- 4:30 pm **Next Generation Data Mining - Data Mining Automation & Realtime-Scoring “on-the-cloud”**  
*Stefan Weingärtner, Dymatrix Consulting Group, Mike Zeller, Zementis*
- 5:00 - 7:00 pm **Networking Aperero**
- 5:30 - 6:30 pm **KNIME Partner Training Session**  
*Schrödinger, ChemAxon/Infocom, Chemical Computing Group*
- 7:00 pm **Dinner - The KNIME UGM dinner is kindly sponsored by our Life Science Partner**

# KNIME UGM Agenda



Thursday March 3, 9:00 am – 6:00 pm, Technopark Zurich, Auditorium

8:00 am            **Registration and Breakfast**

## **Session 3: KNIME Life Science**

9:00 am            **KNIME as a platform for distributed molecular property predictions**

*Nils Weskamp, Boehringer Ingelheim*

9:30 am            **A stitch in KNIME saves nine: Strategies for design in medicinal chemistry**

*Michael Bodkin, Eli Lilly*

10:00 am           **Primetime for KNIME: Towards an integrated analysis and visualization environment for RNAi screening data**

*Oliver Gathmann, Cenix BioSciences*

10:30 am           **Coffee Break**

11:00 am           **KNIME: Applications in Structure-Based Drug Design at Evotec**

*Michael Mazanetz, Evotec*

11:30 am           **KNIME and Next Generation Sequencing: From data cleansing to systems biology**

*Bernd Jagla, Institute Pasteur*

12:00 pm           **KNIME workflow for chemoinformatics tasks at Angelini**

*Candida Manelfi, Angelini Research Center*

12:30 pm           **Lunch Break**

# KNIME UGM Agenda

## Session 4: KNIME Partner Session

- 1:30 pm      **KNIME Life Science Partners**  
*Schrödinger, CambridgeSoft, Tripos, Symyx, Infocom/ChemAxon, BioSolveIT, Cresset, Chemical Computing Group, Cloud Broker*
- 4:00 pm      **Coffee Break**

## Session 5: KNIME Community Session

- 4:30 pm      **The RDKit: open-source cheminformatics (now for KNIME too!)**  
*Greg Landrum, Novartis NIBR*
- 5:00 pm      **KNIME Nodes - Precompetitive Computational Medicinal Chemistry**  
*Michael Bodkin, Eli Lilly*
- 5:30 pm      **Scripting without Scripts: A User-Friendly Integration of R, Python, Matlab and Groovy into KNIME**  
*Felix Meyenhofer, MPI Dresden*
- 6:00 pm      **End of day**

# KNIME UGM Agenda

## KNIME Special Interest Group High Content Analytics

Friday March 4, 9:00 am – 4:00 pm, Technopark Zurich, Auditorium

Note: This is a Preliminary

### Session 2: KNIME HCA Extensions

8:00 am **Registration**  
 9:00 am **Opening**

11:30 am **Overview New Image Processing Plugin**  
 Martin Horn, KoRS-CB University of Konstanz

**Screen Mining with KNIME**  
 Martin Stöter, MPI Dresden

### Session 1: HCA applications

9:10 am **Image Analysis**  
 Martin Stöter, MPI Dresden

**KNIME Bioinformatics Extensions**  
 Karol Kozak, ETH Zurich

**Automated Chromatin**  
 12:30 pm **Lunch Break**

### Session 3: Informatics models for HCA

**siRNA**  
 1:30 pm **High Content Imaging and Flow Cytometry with Phaedra: can we integrate with KNIME?**  
 Frans Cornelissen, TIBBE, Johnson & Johnson

**Challenges in Natural**  
**High-content screening and data pipelines**  
 Jürgen Reymann, Bioquant

**Host in N And**  
**Columbus: Introducing universal integrated image data management and analysis for High Content Screening**  
 Martin Daffertshofer, PerkinElmer

11:00 am **Molecular**  
**iBRAIN: single-cell analysis of high-throughput screens**  
 Berend Snijder, Uni Zurich

3:00 pm **Discussion & Final Remarks**

4:00 pm **End of day**



Thank You.

Mark Your Calendars:

KNIME's 5<sup>th</sup> Anniversary!  
KNIME 1.0 - July, 28, 2006