

Data Literacy Programs that Actually Work

How three global companies use KNIME to empower data citizens to work independently with data, and save hundreds of hours.

Three data literacy programs that achieved measurable results:

- **30 hours saved per quarter** by automating web crawling and creating interactive data analytic dashboards with KNIME and Tableau.

- **Pilot projects were reduced by 80%** through automating project management tasks such as deadline tracking, Kanban boards, and supply chain warnings.

- **Nearly 60 audit use cases** using data analytics in three years after training auditors to use KNIME independently in an internal, tailored program.

The Siemens logo is displayed in a bold, teal, sans-serif font. It is positioned in the upper left corner of the page, which features a background image of a modern industrial factory floor with robotic arms and car chassis.

Driving a Citizen Data Scientist Approach

Getting 3,500 + Data Citizens Working Better and Independently

The demand for automating day-to-day procedures is growing daily. On top of that, billions of bytes of data, multiple data sources, and hours of manual work put into sorting it all out, make these procedures hugely complex and time consuming.

Since January 2018, the Data Visions Team at Siemens has been developing analytical 'products' to support the Digital Industries (DI) strategy and drive a data citizen approach for future collaboration - preparing for the time when data science will eventually become a commodity. The team has made it possible for not only data scientists to work with data, but also data citizens – those charged with pulling insights out of data – to make decisions and drive change.

Today at Siemens over 3,500 data citizens globally, are working better and more independently with data using KNIME Analytics Platform and other integrated software. KNIME has also become an invaluable tool for Robotic Process Automation (RPA), by completely automating many mundane, manual tasks. This has freed up lots of time to work on other areas of the business.

Example: Automating Competitive Analyses with Text Mining and Web Crawling

A project that the Data Visions team got involved in, was automating competitive analyses using financial statements from the internet. The projects' objectives were to search for financial reports on company websites, extract relevant financial statements out of the PDF reports, and transform the statements into a structured format to prepare an internal report. From a business perspective, the requirements were to:

- Automate the extraction of financial reports from the web using a crawler
- Identify key KPIs using template-based text analysis
- Integrate the data flow into an existing report workflow

Saving 30 Hours Per Quarter

Previously, the user did a manual website check of the competition with (often) incomplete information. The repetitive task of analyzing competitors was repeated every quarter in a manual and time-consuming process: search competitor website for PDF, extract relevant information, prepare finance report, send to management. The Data Visions team built a process using KNIME Analytics Platform to automate this entire process. A KNIME workflow, built by a team of data scientists, crawls the competitor's website and downloads financial

Company

Siemens Digital Industries (DI) is an innovation leader in automation and digitalization. Closely collaborating with partners and customers, DI drives the digital transformation in the process and discrete industries.

Solution highlights

30

hours saved per quarter.

Automation & Digitalization

Data Literacy Manufacturing

reports as PDFs. This step is repeated automatically across different websites.

The same KNIME workflow then applies data mining, classifies document types, and extracts values out of the PDF. Then, using the KNIME Integration with Tableau, a report is created and published in an interactive dashboard. Here, the end user can simply open the dashboard and view the results. KNIME has saved this person approximately 30 hours every quarter. That time is now spent on valuable tasks such as analyzing and interpreting the aggregated statements.

Empowering Users to Work Independently with KNIME

As people are the most important factor for success in the digital age, an important aim of the Data Visions initiative is the empowerment of business users to work independently with tools such as KNIME. The advantage of the program is that users can easily start their own projects and are more flexible with what they create. The aim of the coaching approach is to find people that are open minded, enthusiastic regarding data, and willing to learn new concepts. Those people then act as multipliers in their department and show colleagues how to get started with KNIME. Additionally, with the analysis of web data with respect to competitor relevant information, Siemens has been able to introduce more advanced AI-driven topics that generate (with some effort) additional value for the business.

Paving a Way for the Future Through Knowledge Sharing and Transfer

The work that the Data Visions team is doing, resonates well with the entire company. The interest in automating daily work, freeing up time, and empowering everyone to get insights out of data continues to rise. The team has plans to run more internal workshops to share knowledge across different organizations. An internal intranet has been built for everyone to share their KNIME use cases and tips and tricks, as well as a library of how-to tutorials. Meetups offer teams and departments the chance to learn more about KNIME, share use cases and experiences, and build internal communities to collaborate on future projects.

Why KNIME?

The open source and free KNIME Analytics Platform was an extremely easy entry point, making it quick and easy to get started. With one simple download, KNIME is installed on the desktop and data scientists can start building their workflows. KNIME has many useful features, which, compared to other offerings on the market, made it an obvious choice.

The no-coding, graphical UI makes KNIME very easy for a non-programmer to use and the self-documenting features ensure that work remains transparent and easy to follow-up on if someone leaves the team – which was previously a considerable pain point. It's extremely open with so many nodes available that almost anything is possible. And, if there's no node available, then it's possible to develop a KNIME node. The community behind KNIME – both the open source

"The open source, free KNIME Analytics Platform was an extremely easy entry point, making it simple and fast to get started. With one simple download, the platform is installed on the desktop and data scientists can start building their workflows."

Matthias Stephan, Head of Data Visions,
Siemens Digital Industries

community and the KNIME team who are active on KNIME Forum and KNIME Hub – provide fast, valuable support.

Trusted third party extensions provide assurance that there are no risks when working in production, which provides strong peace of mind. The speed of execution has reduced manual effort of everyday work by automating processes. It's now possible to get results in minutes rather than hours! Lastly, KNIME has made it possible (and easier) for collaboration across teams. KNIME components, for example, will play a more important role over time.

The next step for Siemens is to get a KNIME Server license to provide even more functionality around collaboration, automation, and productionization.



Sparking Data Literacy with KNIME and Making Better Decisions

Empowering Business Users Worldwide

Continental is committed to increasing the data literacy of the entire organization by empowering users in business functions (with no IT or programming background) to work more efficiently and effectively with data and, ultimately, discover new insights with that data. Microsoft Excel was previously the go-to tool for business users handling data at Continental. Today, it's perfectly enriched by KNIME Analytics Platform for standard workflows, bigger data volumes, or data discovery. For example, repetitive tasks, such as certain forecasting processes, are now automated in KNIME Analytics Platform. Until recently, this was either done via complex spreadsheets or Visual Basic for Applications Macros (VBA, the programming language built into Excel). However, Continental began experiencing maintainability and performance issues for medium-sized data volumes. There are thousands of Excel users at Continental, all working in diverse business functions: controlling, tax, HR, R&D, production planning, etc. The need to give these users a tool which can automate mundane data processing and reporting tasks resulted in a company-wide initiative to roll out KNIME Analytics Platform.

A Global, Low Barrier, Roll-Out

The roll-out of KNIME at Continental started out as a pilot project in Chassis & Safety (C&S), a €9.6bn sales division of the corporation, and was led by Dr. Arne Beckhaus - Head of Big Data and Digital Transformation C&S at Continental. The very first use case emerged in the Brake System R&D Department, where KNIME was used to generate Kanban boards based on data exported from the issue tracking system.

An agile setup of starting small and expanding quickly had a positive response. It was clear just how much business users loved using KNIME, and how much data processing power was gained. The roll-out was made easier by positive word of mouth and a high, bottom-up demand for trainings on how to use KNIME.

Many processes across the entire organization have now achieved higher levels of automation. For example, production planning, reporting (controlling, logistics, etc.), various project management tasks such as Kanban boards and deadline tracking, supply chain warnings, and so on. Data integration and data wrangling are now mostly done by business

Company

Continental develops pioneering technologies & services for sustainable & connected mobility of people and their goods. Founded in 1871, they offer safe, efficient, intelligent, & affordable solutions for vehicles, machines, traffic & transportation.

Solution highlights

80%

time savings in our pilot projects.

Automotive

Manufacturing

Data Literacy

users in KNIME, and, in many cases, this step alone yields the promised data insight or process automation. Sometimes KNIME is used for an intermediate data preparation step before the data can be exported to visualization tools such as PowerBI, Tableau, and MicroStrategy. Continental also uses Amazon Web Services for long-running workflows (or workflows which are RAM or CPU hungry).

Better Decisions, More Insights, Huge Resource Savings

Significant time savings
of over

80 %

in pilot projects.

Lead time education
from 2 days to

30 min

of month-end controlling tasks.

Run an entire plant budgeting
process at a

10 fold

precision level.

"Since rolling out KNIME at Continental, we're making better decisions, everywhere. We're also able to back these decisions on stronger data, faster. We're gaining more insights from data, and freeing human capacity from mundane data processing and reporting tasks."

Dr. Arne Beckhaus - Head of Big Data and Digital Transformation C&S at Continental

And there's a lot more potential for using KNIME throughout the organization. "One indicator for the high adoption of KNIME is the rising demand for in-house trainings" says Arne. "It's great to see that we can serve this huge bottom-up momentum with a very small central team". The next steps will entail creating more user-specific training content for users in different departments, enabling a more institutionalized roll-out, and realizing and standardizing knowledge sharing across teams and departments. Due to high interest by external parties who want to learn from Continental's data transformation journey, the decision was even made to offer data services to external clients using KNIME. Continental Engineering Services is now an official KNIME Partner and able to offer these services.

Why KNIME?

KNIME was chosen primarily for its business user friendliness but also due to speed, data volume, breadth of functionality, and state-of-the-art data wrangling and data science features.

"From a business perspective, the open source model was the only commercially feasible option to rapidly spread data analytics competence throughout the organization".

Dr. Arne Beckhaus - Head of Big Data and Digital Transformation C&S at Continental

A long tail of infrequent users would have made the license fees too expensive and a flat, company-wide license would have slowed everything down due to the initial investment. Typical commercial business models simply would not have enabled such rapid dissemination. KNIME Server is now also used on selected projects at Continental, enabling teams to collaborate on KNIME workflows and share amongst colleagues.

KNIME is easy to maintain long term and can be enhanced with custom extensions, which are built and developed by Continental. Examples include extensions to enable custom formatting in Excel reports, or to parse third party PDF documents with text position information for automatically analyzing contract information. From an IT perspective, KNIME is easily installed on client PCs worldwide. Continental can pre-package a customized installation and deliver it as a standalone software application - making it easy for KNIME to be included in the official software distribution process.

Lastly, rolling out KNIME to business departments has helped raise the awareness of analytics and AI at the management level. They now understand the advantage of finding answers to problems by effective data wrangling and the necessity of good quality data for applying data science algorithms in KNIME. The result? Increased data literacy across the organization, more efficient, automated processes, new insights from data without blind application of AI, and better decisions everywhere.



Rabobank

Anomaly Detection: Unknown Patterns in Anti-Money Laundering

Money laundering makes large amounts of money that have been obtained through illegal activity appear as if they have come from a legitimate source, and avoiding this is a top priority for auditing teams. The key element in detecting money laundering is understanding known vs unknown patterns.

Senior auditors at Rabobank are first interviewed to understand what identifies the behavior of people with wrong intentions. Those hypotheses are then translated into a set of defined business rules. KNIME helps significantly by defining these rules and applying them to the entire portfolio, which is one of the biggest wins as it increases quality assurance. Because instead of a sample of twenty, it's possible to test, approve, and show across an entire population to uncover known patterns.

However, in money laundering, unknown patterns are more interesting. The chances of missing something are also much higher because individuals with bad intentions quickly learn how to avoid patterns that are already known. Anomaly detection, cluster analysis, and text mining are a few ways to find unknown patterns. With machine learning, it's possible to learn what the currently unknown patterns are.

Empowering 300 Auditors to Run Anomaly Detection Independently

This machine learning approach was applied using KNIME and needed to be made available to all 300 auditors - most of whom were non-tech. A KNIME workflow was developed, which is reusable and easy to handle for those who have even just a small amount of KNIME knowledge. The basic steps include accessing and inspecting the data, conducting anomaly detection, creating a visualization, and exporting and storing the output. The Data Explorer node enables users to first inspect the data and better understand it, which, in the case of anomaly detection, is essential. Under the hood of the basic KNIME workflow, which is largely made up of KNIME components, are many more KNIME nodes. This approach is valuable because it hides the data science complexity, but still makes it possible to adapt, reuse, and share the workflow as needed.

Company

Rabobank is a Dutch multinational banking and financial services company. It is a global leader in food and agriculture financing and sustainability-oriented banking.

Financial Services

Audit

Anomaly Detection

Data Literacy

How Rabobank is Building a Data-Driven Culture in Audit

Rabobank management recognizes that the digital world is changing rapidly and is becoming more and more data driven. This is having a profound impact on the auditing profession, auditing services, and on the auditors themselves. Three data-driven targets were therefore set, and the message from management was clear: just get started!

1. Become more efficient with data analytics
2. Apply more data analytics in the audits that are being done
3. Have datasets readily available for deep dives, follow-ups, and Q&A

About 90% of what Rabobank does is basic analytics, which is already creating a significant impact with, for example, descriptive analytics, dashboarding, data quality profiling and reconciliation, benchmarking, as well as translating knowledge of senior auditors into business rules for testing. The remaining 10% is more advanced. For example, trend analysis, process mining, anomaly detection, cluster analysis, and text analysis.

It took approximately one year to understand that it's not just about hiring a data scientist to do data analytics. The Audit team therefore uses a target operating model to embed their own way of working with data. Rabobank adopted an innovation cycle, and in two days all 300 auditors were trained to become data literate and to start using KNIME themselves. Whilst it was recognized that they might not be the ones ultimately doing the analysis, emphasis was placed on the importance of being able to see and understand the potential in the data and how it can be used. Through this, the culture has become more and more data driven. The team's unique technical architecture includes KNIME as one of the key tools for data analysis.



Results

The approach at Rabobank has been to start small and scale fast. To test the value of data in audit the first year consisted of seven use cases. In the second year this increased to 13, in the third year 35, and in 2020 there were 57 auditing use cases taking advantage of data analytics - which was only possible via empowering the auditors to do their own analyses.

Next Steps

There are many ways to get started with KNIME - from downloading and trying it out, to attending an event and seeing it in action:

Download KNIME



KNIME Analytics Platform is the free and open source software for creating data science.

www.knime.com/downloads

Visit KNIME Hub



Thousands of blueprint solutions are available for getting started on your data projects right away.

www.hub.knime.com

Join an Event



Webinars and courses are happening on a weekly basis, plus regular Data Talks to extend your knowledge.

www.knime.com/events

If you need advice on tackling your next data project, or tips on making KNIME the go-to analytics tool in your organization, our Customer Care team is happy to help.

Reach them here: www.knime.com/contact

KNIME Software

KNIME Software consists of **KNIME Analytics Platform**, which is the free and open source software for creating data science. It's intuitive, open, and continuously integrating new developments to make understanding data and designing data science workflows and reusable components accessible to everyone. **KNIME Server** is the enterprise software for team-based collaboration, automation, management, and deployment of data science workflows as analytical applications and services. Non experts are given access to data science via KNIME WebPortal or can use REST APIs.



SIEMENS

Continental

 **Rabobank**