

“Data science starts in school.”



When Ralf Klinkenberg founded an open source project in 2001 that was the precursor to the company RapidMiner, the value of data was not yet clearly understood. Since then, not only his company, but also the handling of data has changed rapidly.

COPY — Sabrina Waffenschmidt

His job is to look into the future. And this vantage point is more in demand than ever. Ralf Klinkenberg is co-founder and Head of Data Science Research at RapidMiner, one of the world's most popular software platforms for data science, data mining, and predictive analytics. Unmistakably like no other, it helps companies to generate forecasts from their data.

Klinkenberg has developed the software since 2001 together with Dr. Ing. Ingo Mierswa at the Technical University of Dortmund as a flexible open source tool that examines large amounts of data for trends and associations and facilitates organizations' internal work with data mining. Whether Lufthansa, Intel, or BMW, PayPal, Ebay, or Siemens, with each new customer the two founders expanded the user base of their software. Today, RapidMiner has more than 380,000 registered users in over 100 countries.

MORE THAN 100 PARTNER COMPANIES

While RapidMiner initially followed the relatively typical service-based business model of an IT service provider, today the company relies on a licensing model with more than 100 partners worldwide. "In addition to being innovative, IT organizations need to build large ecosystems and communities of developers to survive. Companies that believe they can do it all alone will perish."

Klinkenberg and his business partner recognized this early on. They measure their success mainly based on downloads and subscription rates, so they can easily understand what works and what doesn't. "In the beginning, RapidMiner was a tool for experts. The first version had no design interface and was difficult for new users to utilize. When we added the graphical environment, user data rose dramatically. Similar to when we added the installer."

Today, the tool offers online tutorials, application templates, and transparent example processes with sample data records and an auto modeler function. The most important thing for Klinkenberg: "Facilitate operation and increase automation while maintaining transparency and flexibility. Add to that our large community and marketplace, so other companies and universities can build their own projects based on our software and share RapidMiner expansions with the community."

DATA SAVE LIVES

RapidMiner now works across all industries: Automotive, aviation, chemical, metalworking, food, insurance, banking, internet. Above all, the topic of 'predictive' is now established and growing rapidly. Nevertheless: Many organizations are still in the testing phase. Even industries and companies that are already engaged in predictive analysis

still have much untapped potential, says Klinkenberg. "Although many machines are already connected in industry, there are still many unused interfaces along the production chain. And the chain itself is often forgotten."

According to Klinkenberg's assessment, the greatest effects can be achieved in production and industrial manufacturing through the comprehensive use of data: Production becomes more plannable and efficient, products are better and more individual, and environmental resources are conserved. Klinkenberg also sees strong effects in the healthcare and medical industries. Proper use of data will not only allow for more personalized treatments, it will also help prevent medical problems such as heart attacks and strokes – and with the appropriate treatment recommendations.

BREAK UP DATA SILOS

To work efficiently with data, it is particularly important to break up data silos and to connect the data across all processes. Klinkenberg advises: "Just start somewhere. Often you can achieve a lot with minimal effort. But most of all, you acquire a sense of your own data quality along the way. External consulting can help, but it is becoming increasingly important for companies to build their own data science competence. This creates a real competitive advantage."

BUILDING SYSTEMATICALLY AND WITH CARE

Since 2013, the company has also maintained a location in Boston and thus has a direct comparison of developments between the US and Germany. "The US is a bit ahead of Germany and Western Europe. Americans are often quick to opt for an innovative solution but throw it overboard more quickly if it does not work right away. In Germany, the start-up time is a bit longer, but development is more systematic and sustainable."

Politicians must also become more aware of data science, demands Klinkenberg. "It is important to me to establish the topic in Germany more strongly in politics and to reduce fears in society. Germany is a very strong industrial nation and on a good trajectory. But the wheel is turning faster and faster, and local companies are now competing with global corporations. There are huge potentials and they have to be recognized. Certainly not everyone has to become a data scientist, but everyone has to be aware of it. This should already be taught in school."



Vita

After studying computer science at the Technical University of Dortmund and the Missouri Institute of Science and Technology (MST) in the USA, Ralf Klinkenberg began researching machine learning, data mining, text mining, and predictive analytics at both universities in 1996. In 2001, together with Ingo Mierswa, he founded the project and later the company RapidMiner, which Klinkenberg leads today as the Head of Data Science Research.

 rklinkenberg@rapidminer.com

 rapidminer.com

www.t-systems.com/solutions/predictive