

The good, the bad and the ugly

SmartStream's Chief Innovation Officer **Andreas Burner** oversees the company's research into AI that can identify bad and missing data for reconciliation purposes. In insurance, dirty data isn't just a financial burden but, increasingly, an ethical one



In June, the US-based Center for Economic Justice singled out the insurance sector as being particularly unimaginative and slow-moving in its approach to data bias.

"In an era of big data analytics and insurers' rapidly growing use of third-party data and complex algorithms, the potential for algorithmic bias and proxy discrimination has grown dramatically," it said.

A study last October in the journal *Science* claimed an algorithm used by hospitals and insurers in the US miscalculated the care needs of sicker black patients in particular – partly because these patients tend to avoid having to go to hospital to pay for services, due to not having adequate insurance cover.

Separately, the Center for Economic Justice statement quotes a recent paper from scholars at the California Law Review, which illustrates why campaigners are concerned about AI use cases in insurance.

"Data mining can inherit the prejudices

of prior decision-makers or reflect the widespread biases that persist in society at large," they said. "An algorithm is only as good as the data it works with."

As someone dedicated to applying AI to data in 70 of the world's top 100 banks and many other key institutions, including insurers, Andreas Burner is well aware of that fact.

Part of his job as SmartStream's chief innovation officer is to oversee the company's Vienna-based Innovation Lab, where computer scientists, mathematicians and banking experts are collaborating to find use cases for AI in the financial industry. A crucial part of that task is identifying and rectifying the problems that 'bad' data collection and entry can create further down the line. As a company, SmartStream intends to invest 20 per cent of its yearly revenue into research and development in the lab, which opened in 2018 to focus on potential uses for AI, machine learning and blockchain.

"If the data is biased, the machine learns from that," Burner says. "We've seen several examples where applications

were automatically filtered, and it has bad bias in there. Recently, there have been many papers on how to identify that bias, because it's so important to understand and try to remove it from the data."

The papers that Burner references are the first pioneering efforts to shrug off inherited discrimination, ensuring society's inequalities don't piggyback into the age of AI. But, as the Center for Economic Justice points out, progress is simply too slow – and the insurance industry is ill-equipped to act on the recommendations. "AI and machine learning technology is really complicated. You need experts to master it – and it takes years," Burner points out.

Several of the world's leading insurers have already implemented AI to help underwrite, price and risk-assess their policies. AXA, Aviva and Lloyd's of London have invested heavily in the technology in recent years, and disruptors like Cuvva and Lemonade Insurance have based their entire business models around it, with chatbots at the front end and AI underwriters in the back office.

This innovation should mean more competitive prices for consumers, yet lingering over this is the concern that the cost reductions associated with the rise of AI may not be equally distributed throughout society. As talk of AI regulation builds, it may fall to fintech specialists in data hygiene and processing to provide ways to dissect and nullify bias in the data AI is founded upon.

In his own sphere, Burner hints that SmartStream is building such a capability – albeit with individual financial institutions, addressing their specific data discrepancies.

"We are running proof of concepts in the innovation lab at the moment where we have huge amounts of data to help an institution that is having problems with the quality of its data," says Burner. "Either data is missing or brokers deliver bad quality data. We have implemented one of our machine learning libraries to retrieve that data from other systems and try to identify everything that is missing. You cannot imagine how successful that is – we'll be writing a white paper on it because it's really a story to tell."

It could be one that others can learn from in the wider context of AI use.

SmartStream's lab-based approach has been compared to the workings of an F1 team, where specific, high-tech solutions are developed by global experts working at the cutting edge, which then filter back into the financial services industry. For SmartStream, every eureka moment in the Vienna lab has implications for the handling of data throughout the financial sector; every data hygiene breakthrough makes AI more thoughtful and more conscious of the world beyond the data that it devours.

And Vienna continues to deliver. Most recently, at last year's Sibos event in London, the lab unveiled SmartStream Air (AI in Reconciliations), the result of 18 months' rigorous work, using real bank data to ensure the efficacy of the solution. SmartStream's new reconciliations software, Cloud-based and simple to onboard to, created a huge buzz. Banks

and insurers found the software delivered solutions to data inconsistencies, disputes and exceptions in minutes, rather than weeks. And like other Cloud tech providers, SmartStream has found its solution particularly popular in the last few months.

"Now we are in this COVID-19 lockdown, it's quite tough for organisations to purchase and install hardware and software, because people are just not on site," says Burner.

"So we've seen an increase in Cloud technology recently. In these times, it's actually the optimal solution for organisations that want to quickly and easily verify their data."

SmartStream AIR proves that the work in Vienna can have profound industry-wide effects – helping all manner of financial institutions cope with their swollen reservoirs of data.

"Insurers also have lots of data within their organisations, in different databases

“If the data is biased, the machine learns from that... it's so important to understand and to try to remove that bias

or data lakes," says Burner. "So SmartStream Air is really a tool that can be used in any organisation that wants to perform data verification and identify problems, so that they are resolved quickly."

AI in reconciliations is ultimately about back-office savings. Significant ones. Those complex, interweaving data trails once struggled over by

manual data workers are child's play to SmartStream AIR. Insurers are saving hundreds of hours and countless frayed nerves using SmartStream's software, and the reconciliations AIR produces facilitate straight-through processing of datasets across insurers' verticals. With better-shaped data, insurance firms can concentrate on their product lines and the processes upon which they rely.

"That pre-process, making the data

better, has such a positive effect," Burner says, "because everything from then on is straight-through: it's being processed automatically, with less manual touch points for users, and much more quality for managers. It's really, really impressive."

SmartStream AIR was created after countless conversations with SmartStream's customers, which directed the Vienna lab towards automated reconciliations as a priority for financial institutions. But it can be used to reconcile not just financial digits. There are applications across an organisation's departments.

"We also see that HR, for example, have need of such tools. And big organisations typically hold customer names and addresses in their CRM as well as other systems. Whenever you want to synchronise data quickly and easily, SmartStream Air is a perfect choice for that," says Burner.

Given the recent calls for closer scrutiny of AI in insurance underwriting, SmartStream could perhaps help insurers identify ethical concerns as much as efficiency-saving ones. Because there's no doubt that AIR is a tool with big brain. Burner jokes that it can solve 3.5 million arrays in a Sudoku puzzle.

"But, what's more important is, if the person behind the sudoku made errors, SmartStream Air can identify those. It will be automatically reconciling the Sudoku, while solving it, and tell the creator where he or she made typos.

"That's actually what reconciliation does. It identifies exceptions and disputes, in huge amounts of data."

Bad data dilemma: The SmartStream lab is working to eradicate it