

Case Study

Llaama Helps Biopharma Companies Bring New Treatments to Market Faster with Kalix



Llaama is a European start-up dedicated to enabling full traceability and reproducibility in complex AI and data science workflows. With its key solution, T2R2, Llaama aims to bridge the gap between research and development, facilitating true translational science capabilities for biopharma companies.

“What Kalix gave us out-of-the-box would have taken months of DevOps and database administration to achieve ourselves. We would rather invest our time in adding direct value related to our core domains of AI and Biopharma.”

Bernard Deffarges,
Co-founder and CEO, Llaama

The Need

Reproducibility is a principle that underpins the scientific method – the wider community accepts experimental results only when they can be reliably reproduced. However, as the complexity of research projects rises, it becomes increasingly difficult both to reproduce results and to understand what has gone wrong if the results are not as expected.

In the biopharma industry, the reproducibility challenge is particularly acute. Most biopharma companies are divided into functional silos – research, development, marketing, and so on – which exacerbates the challenge of trying to translate a positive result in research into an effective and approved clinical practice or drug.

Achieving this translation is a high priority in biopharma: when results are not transparent, traceable, replicable, and reproducible, trying to plug the gaps is difficult, time-consuming, and costly. Llaama estimates that enabling seamless translational science through improved reproducibility could save more than USD 7 billion annually in clinical studies alone.

The Challenge

Bernard Deffarges, Co-founder and CEO at Llaama, says: “We set out to reduce the time and effort to turn positive research results into profitable new drugs and clinical practices by making the translation smooth, transparent and traceable.”

In seeking to enable full traceability and reproducibility in complex AI and data science workflows, Llaama knew it was vital to create a distributed solution that could run anywhere and everywhere, from on-premises to private cloud to edge to public cloud. This is because most biopharma companies have data and applications in all these places, and they also share data and code with external partners such as contract research organizations (CROs).

“Research activities may be fragmented across different internal and external teams, so biopharma companies want a better way to track exactly what is happening at every step and in every location,” says Bernard Deffarges. “Our vision was a solution to instrument all the software containers involved, and use resulting data to build directed acyclic graphs [DAGs]. These DAGs would enable the tracing of every step from the data up through all the branches to the final results.”

From the technology perspective, Llaama wanted a lightweight, non-invasive way to enable container instrumentation and observability across distributed environments.

“The streaming features within Kalix make it really **smooth**, **fast** and **responsive**.”

The Solution

Llaama chose Kalix from Lightbend as the foundational technology for its T2R2 solution.

“What Kalix gave us out-of-the-box would have taken months of DevOps and database administration to achieve ourselves,” says Bernard Deffarges. “We would rather invest our time in adding direct value related to our core domains of AI and Biopharma.”

Kalix is a platform-as-a-service that abstracts away the complexity of event-driven microservices so that teams can focus solely on building business logic. It uses the Akka framework for creating distributed multi-agent systems that scale dynamically.

“We used Kalix to build a distributed system for data analysis and AI pipelines that are traceable, transparent, replicable and reproducible by design, hence the T2R2 name,” says Bernard Deffarges. “Our approach is to create a data lineage not as something built on top of an experimental process, but rather as something that comes from the process itself.”

T2R2 uses Kalix to track every event across distributed sets of containers, creating an immutable extended audit trail that shows the logical provenance of every result. Hash codes are calculated for every input and output across all containers, so that the audit trail is cryptographically guaranteed. The solution includes observability at every level: OS kernel runtime, containers, user space and applications.

“Even when an analytical process is working on a huge data set and is distributed across different machines in different locations, T2R2 tracks exactly what happens and where it happens,” says Bernard Deffarges. “The streaming features within Kalix make it really smooth, fast and responsive.”

T2R2 features a bottom-up AI-powered learning workflow engine, enabling it to deduce workflows based on observed processes and then build DAGs of research and analytical activities. Once validated, workflows can be promoted to templates that can be cloned and reused. In this way, the exact processes used in research can be employed downstream in biopharma companies, making it faster and easier to take new drugs and clinical techniques through regulatory approval and to market.

Llaama is now also using Kalix to enable flexible data contracts within T2R2. These are designed to enable biopharma companies and CROs to agree on data schemas at the outset of a research project and lock

them into a contract. Any change to the data schema must then be formally agreed on by both sides. This avoids a frequent problem: where one party changes a variable or a process without informing the other, causing errors and exceptions further down the line.

“With Kalix, we can use protobuf to easily define a data schema that can evolve over time, but only if both partners agree,” says Bernard Deffarges. “With contracts agreed programmatically between partners, T2R2 helps to avoid disputes.”

The Results

Suzanne Studinger, Chief Quality Officer at Llaama, says: “Traceability is a crucial requirement in pharmaceutical development; biopharma companies need to know who did what, when, and why, even down to the level of the infrastructure. Drawing on features such as event sourcing in Kalix, our T2R2 solution can provide easy and secure data-driven execution for traceability.”

By providing auditable traceability across even the most distributed of architectures, T2R2 supports biopharma companies in achieving rapid translation between clinical research and development.

Bernard Deffarges says: “One of the biggest values of our solution is enabling compliance by design in a highly distributed system. This also contributes strongly to cutting the time to market for new drugs or clinical approaches. Being able to beat competitors to market can translate into huge profits.”

Many quality control process steps in biopharma demand the four-eye review principle; T2R2 saves time and effort by replacing this with programmatically guaranteed data lineage.

“There are solutions that can reproduce a workflow, but with no guarantee that there’s traceability or that the output is going to be always the same,” says Suzanne Studinger. “With Kalix, T2R2 offers reproducibility of both workflows and results.”

Bernard Deffarges concludes: “We have built an innovative, bottom-up solution to the deep challenges of reproducibility in biopharma, drawing on the actor model embodied in Kalix. With Kalix, we can focus on our business, following Domain Driven Design (DDD) using protobuf, while the distribution and clustering of actors, the deployment, and everything else happens like magic! We are now working with early adopters to refine the solution and finding industrial partners to bring it to new markets.”

Learn more at [Kalix.io](https://kalix.io)