

Since 2019, PARADIM has deployed a streaming data architecture for real-time collection, collation, analysis, storage, and access of data from research instrumentation.

In parallel, a DMREF for data driven discovery has developed OpenMSIStream, a software toolkit easing adding of data streaming to commercial instruments.

An important consideration in both of these efforts is how to protect the confidentiality, integrity, and verifiability of streamed data, while ensuring data is accessible at the appropriate time to provide data alongside publications.

KafkaCrypto was **developed by PARADIM** to provide end-to-end encryption and data assurance for streaming materials data. It relies on a core symmetric crypto ratchet implementation to provide forward secrecy, and verifiability of what data was produced and consumed when and by what. Together with a Noise protocol framework based asynchronous key exchange, it provides the data confidentiality, integrity, and verifiability needs of users.

OpenMSIStream now incorporates KafkaCrypto and both are now available to the community for adoption and further development.

M. Eminizer *et al.* [J. of Open Source Software 8, 4986 \(2023\).](#)

Codes: <https://github.com/openmsi/openmsistream>
<https://github.com/paradimdata/kafkacrypto>

Tyrel M. McQueen, Margaret Eminizer, David Elbert (JHU)

1. Open-Source Laboratory Streaming Data
2. Encryption in Producers and Consumers
3. Secure Data between Users and with the Platform

