



FINOS

Fintech
Open Source
Foundation

CASE STUDY

The FINOS Legend
Studio Pilot: an Open
Source Success Story
in Financial Services

**DRIVING EFFICIENCY AND
INTEROPERABILITY THROUGH
SHARED DATA MODELS**

LEGEND QUICK FACTS

Date Open Sourced:
October 2020

Languages:
Typescript, Legend-Pure, Java

Link:
legend.finos.org

GitHub:
github.com/finos/legend

The power of visual data modeling

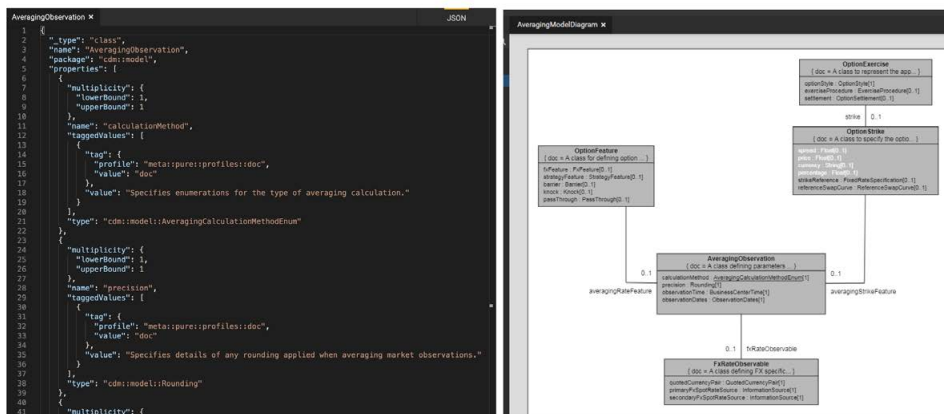
In November 2019, at FINOS' flagship conference, the **Open Source Strategy Forum**, FINOS Platinum member Goldman Sachs announced its intention to open source its internally developed logical modeling language and visual data modeling platform used to build, design, and execute data models.

Legend provides an excellent opportunity to bring subject matter expertise to data models, as its single, visual platform enables engineers and non-engineers - regardless of their technical background - to develop data-centric applications and data-driven insights.

The **Legend language**, together with four Legend platform modules - **Studio**, **SDLC**, **Engine**, and **Shared** - were open sourced into FINOS in October 2020 and are now available for anyone to use.

The launch followed the completion of the Legend Pilot, in which leading investment banks, such as Deutsche Bank, Morgan Stanley and RBC Capital Markets, used a shared version of Legend, hosted on FINOS infrastructure in the public cloud, to build extensions to the Common Domain Model (CDM) developed by the International Swaps and Derivatives Association (ISDA).

Within months of open sourcing Legend, Waters Technology recognized the value and potential of this project by awarding FINOS the American Financial Technology Award 2020 for Best Collaboration initiative.



The Averaging Model developed by the Legend Pilot FX Options Working Group in JSON format (left) and diagram form (right).

What is Legend?

Legend, developed by Goldman Sachs' Engineering team over seven years, provides an end-to-end data platform experience covering the full data lifecycle. It encompasses the Legend-Pure language and a suite of data management and governance tools known as the Legend Platform, which allows engineers and non-engineers alike to describe, connect, visualize, and query high-quality data.

The Legend-PURE **language** is an immutable functional language based on the Unified Modeling Language (UML) and inspired by the Object Constraint Language (OCL).

The Legend Platform is a collection of tools, including the Legend Studio visual data modeling environment, that are used to design, build, and test models in an intuitive, business-friendly way.

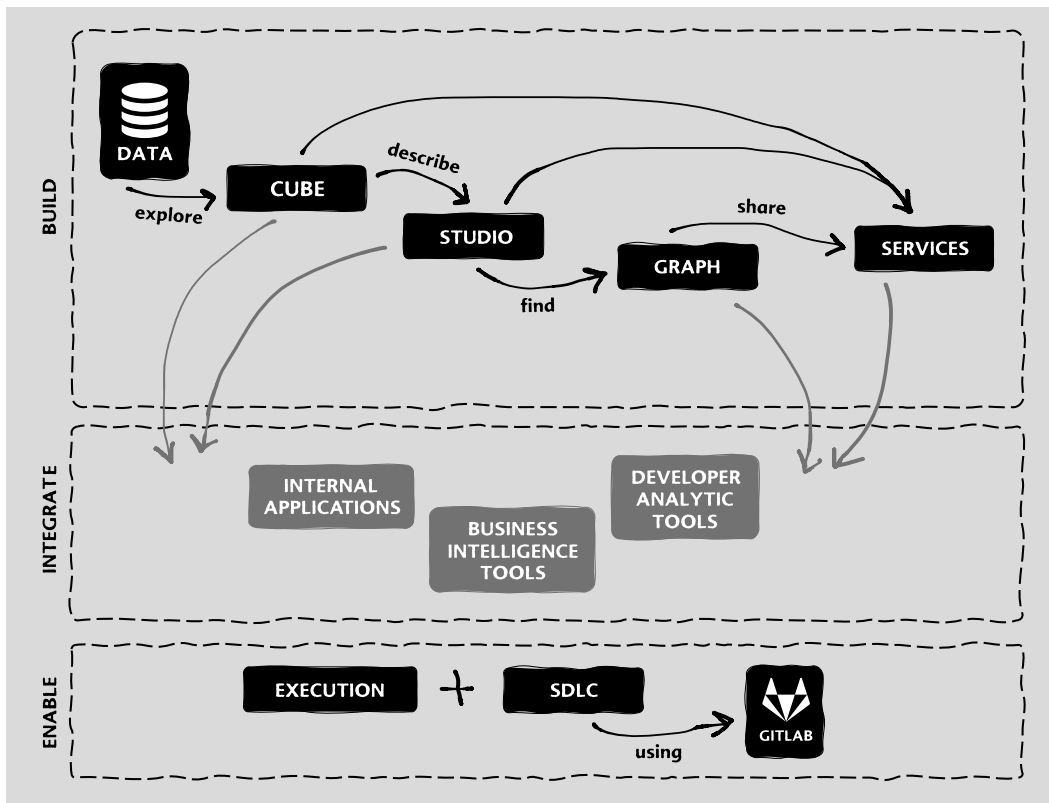
Behind the scenes, Legend is backed by a powerful Execution Engine, and uses GitLab for its Software

Development Lifecycle (SDLC). This enables users to leverage their data models in business workflows, while ensuring that data models are safe to use in live business processes.

Legend supports flat data sources (e.g. csv) and relational execution, and can generate code in languages and formats such as SQL, JSON, Protobuf, AVRO, and Rosetta JSON Schema, which enables easy interoperability across models and systems.

Other Legend components include Cube UI for user-friendly tabular data exploration, Query UI to search and access modelled data, and Services UI, which will allow managing of production data services (APIs) for repeatable and controlled data access. Cube, Query and Services are on the roadmap to be open sourced.

See the Legend roadmap at legend.finos.org.



The FINOS Legend Pilot

Between April and September 2020, Goldman Sachs led several data modelling efforts as part of the FINOS Legend pilot project. The Pilot brought together developers and subject matter experts from financial institutions, technology firms and industry associations across the US, Europe and Latin America, to develop common data models that would enhance existing industry standards.

Gauging the Industry's Appetite for Shared Data Modeling Tools

A key premise of the pilot was to add value back to the industry -not just demonstrate it as a tool - and FINOS provided the ideal environment for the industry to evaluate Legend.

The pilot working groups were led by subject matter experts from Goldman Sachs' Securities Division, who also proposed use cases and helped the FINOS community become familiar with the Legend platform.

Pilot participants focused on two specific use cases to use Legend to build extensions to the ISDA Common Domain Model (CDM), particularly around FX Options and Commodity Reference Data.

At the same time, Goldman's Engineering team worked on two technical fronts: they worked closely

with Regnosys - curator of the [Rosetta DSL](#) used in the ISDA CDM - to translate the definitions of the ISDA CDM into Legend-Pure, and in parallel they abstracted and modularized the Legend code to prepare for its open sourcing.

Enabling Open Source Collaboration in a Highly Regulated Industry

FINOS enabled open collaboration through three different, but complementary, pillars: governance and policies; technical support; and community engagement.

The Foundation's **governance and policies** ensured an independent, neutral environment for pilot participants to collaborate in a safe and compliant way, something particularly important for a highly regulated and competitive industry like financial services. FINOS executed and handled the licensing agreements with modelers participating in the Legend pilot and enabled the necessary tooling to ensure compliance.

On the **technical front**, the Legend project collaboration leveraged tools and services developed by the FINOS [Open Developer Platform](#), providing a secure and compliant collaboration and development workflow. Specifically, the FINOS Infrastructure team deployed an [external GitLab instance of Legend](#) on AWS for pilot modeling activities, created a [GitHub repository](#) with its corresponding tooling, and provided the necessary meeting infrastructure.

FINOS tapped into its **diverse community** of developers and subject matter experts across the industry to convene a critical mass of modelers to pilot Legend. The FINOS team supported the pilot project with project management and project marketing activities and managed operations, including modeler onboarding, meeting scheduling and communications.



From our participation in the FINOS pilot, we believe that Legend Studio holds promise to enhance collaborative and federated data architecture and modeling within the bank and the industry.”

- **RUSSELL GREEN - HEAD OF CLOUD ARCHITECTURE, DEUTSCHE BANK, AND FINOS BOARD MEMBER.**

“ We’ve witnessed firsthand the struggle with data silos, duplication and quality as the complexity of data grows dramatically. And we’ve heard the same from peers in the industry. Our data platform, Legend, starts by breaking down silos and building a critical bridge over the historical divide between business and engineering, allowing companies to build data-driven applications. We’ve now made this platform available for anyone to use as open source code and are thrilled to continue collaborating with our peers and other industry participants under the FINOS umbrella to drive industry-wide data & model standards.”

- **PIERRE DE BELEN, HEAD ARCHITECT OF THE LEGEND PLATFORM,
MANAGING DIRECTOR & TECHNOLOGY FELLOW, GOLDMAN SACHS**

Achieving Increased Efficiency Through Standardized Data Models

ISDA, the International Swaps and Derivatives Association, made its Common Domain Model (ISDA CDM™) available to Legend pilot participants through the FINOS hosted Legend instance.

The ISDA CDM

The ISDA CDM is a blueprint for how derivatives are traded and managed across the trade lifecycle, and it aims to serve as the single, common digital representation of derivatives trade events and actions.

With its hierarchical representation across trades, portfolios and events, the ISDA CDM promotes transparency and alignment in financial markets and allows for the production of consistent trade data, in turn enhancing risk management and trade processing capabilities for industry participants.

Importantly, the CDM - which uses the Rosetta DSL - is available in machine-readable and machine executable formats and languages that can be consumed by various technologies, which reinforced its suitability for the FINOS Legend Pilot Project.

Source: [ISDA CDM Factsheet](#)

Over the course of the Legend Pilot, the ISDA Development team guided pilot participants through the CDM and contributed to the extensions to the CDM. ISDA worked closely with Goldman Sachs’ Engineering team and Regnosys to translate the CDM model into the Legend-Pure language.

Standards: overcoming data inconsistencies and fragmentation with Legend

Challenges to data management in finance are less related to the volume of data processing, than they are to the fragmentation of data and the great number of actors participating in the processing flow, which creates communication issues, processing breaks, and other inefficiencies.

Legend provides a solution to these challenges by focusing on data discovery, helping technical and non-technical users to find data; data transformation, linking information together and building a leverageable lineage graph; data quality, improving quality checks all along the processing pipelines; and data delivery, enabling all actors to easily and safely acquire data.

Now that it’s open sourced, Legend makes it possible for industry participants to come together and describe data in a consistent way, facilitating the definition and adoption of common, industry-wide data standards.

Bringing the Industry Together

Along with Goldman Sachs, FINOS members Deutsche Bank, Itau Unibanco, Morgan Stanley, RBC Capital Markets, ScottLogic, Wells Fargo as well as Digital Asset, ISDA and REGnosys participated in the pilot. The community’s contribution was key to building shared data models, identifying use cases, and providing feedback on Legend features and functionality.

An Open Source Collaboration Success

The Legend pilot showcases the value and potential of open source collaboration, software and standards in the financial services industry. Within months of open sourcing Legend, Waters Technology recognized the value and potential of this project by awarding FINOS the American Financial Technology Award 2020 for Best Collaboration initiative. The award recognizes the collaborative work conducted on the open-sourced Legend platform, done in conjunction with Goldman Sachs, Morgan Stanley, Deutsche Bank, and RBC.

As part of the FINOS Legend Pilot, participants in the FX Options and Commodity Reference Data working groups built extensions to ISDA’s CDM that were proposed back to ISDA’s Architecture Review Committee in the Fall of 2020.

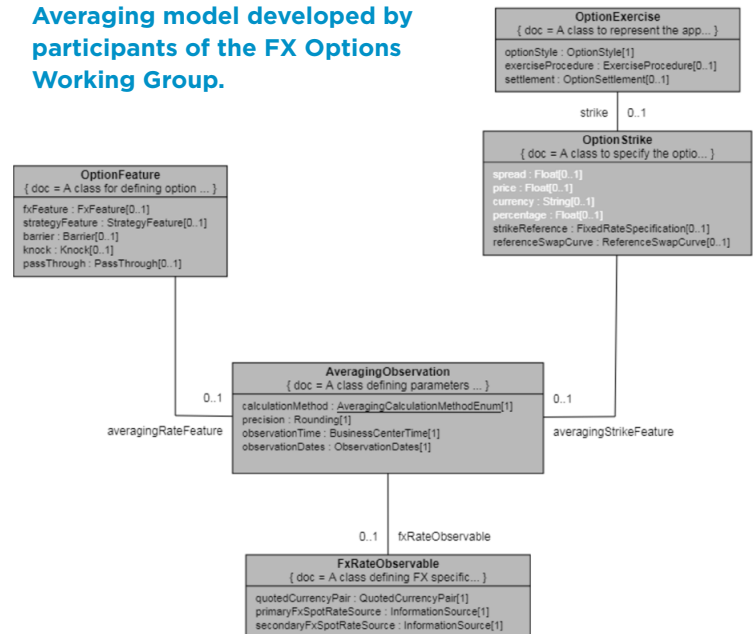
The Averaging Model - an extension to the ISDA CDM FX Options model - built by the FX Options working group was approved and **released in the CDM (see 2.82.8)**.



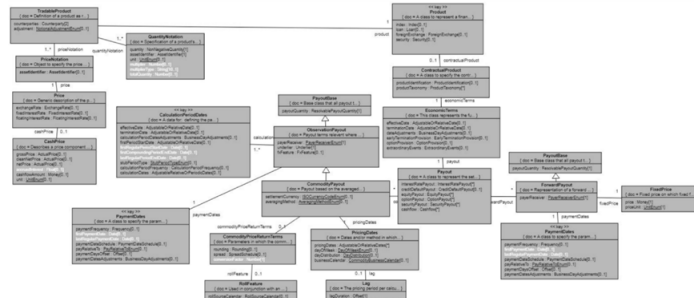
Legend is an impressive technology with great potential for improving industry efficiency. We see tremendous potential for synergies between Legend and our own Morphir project, also open-sourced through FINOS, just last month. **These contributions together validate the FINOS model of bringing industry competitors together to solve industry challenges.**

- STEPHEN GOLDBAUM,
EXECUTIVE DIRECTOR,
MORGAN STANLEY

Averaging model developed by participants of the FX Options Working Group.



The Commodities Reference Data Working Group created a Commodities Swap model in the ISDA CDM which was approved by the ISDA Architecture Review Committee in December 2020 and was **released in the CDM (see 2.99.5)**.



Commodity Payout created by participants of the Commodity Reference Data Working Group

Open collaboration on the Legend Pilot meant that concerns and ideas from different industry participants were addressed early on in the process, which accelerated the review and approval cycle of the proposed extensions. While new extension proposals usually undergo three to six iterations before being approved by the ISDA Architecture Review Committee, the FX Options Averaging model was approved without iterations.

A Win-Win for All Concerned

The FINOS Legend pilot overperformed on its goal to assess the value of Legend for the financial services industry: it created significant value for individual participants and for the industry as a whole, and established a successful model for future industry-wide collaboration.

Individual participants were able to accelerate their journey towards open source readiness

A spotlight on derivatives contracts

Derivatives contracts evolve frequently, have deeply complex, tangled structures, and are processed by many different parties within and outside of each firm. This means that derivatives contracts leverage different schemas that require transformations. As many of the processing steps emerge at different times, these schemas are built with different technologies, require different query engine capabilities, and many validations to encode the requirements of all actors.

Legend addresses these challenges by:

- letting users express models, in a central environment, that can be converted to different technologies in a central environment
- letting users define mappings between different models
- enabling users to define constraints that can be enforced anywhere in the workflow
- operating all these models and transformations in a strong software development lifecycle (SDLC), which optimizes communication among different participants, and accelerates potential modifications and improvements of the financial terms

by deepening their knowledge of open source collaboration tools and best practices in general.

Importantly, non-engineer participants leveraged Legend's visual modeling feature to develop extensions to ISDA's CDM.

“ **The Legend pilot proved FINOS to be a great vehicle for open source enablement of the CDM for ISDA and its members.**

Community use of the shared environment to do even more collaborative modeling will generate more interest in Legend, as well as the CDM and other standards maintained by ISDA. This is a win-win for all concerned.”

- IAN SLOYAN,
DIRECTOR,
MARKET
INFRASTRUCTURE
AND TECHNOLOGY, ISDA

“ Legend provides both engineers and non-engineers a single platform that allows everyone at Goldman Sachs to develop data-centric applications and data-driven insights.”

- ATTE LAHTIRANTA
CHIEF TECHNOLOGY
OFFICER,
GOLDMAN SACHS

On a broader note, it has shown that the Legend value proposition is not intrinsic to internal Goldman Sachs use cases, but rather the opposite: Legend enables engineers and non-engineers to develop shared data models, addressing data efficiency and governance challenges that are pervasive throughout the industry.

The Legend Pilot has established a model for future collaboration, where FINOS:

- provides a secure environment for the financial services community to seamlessly develop open source code and standards,
- financial services institutions contribute use cases and code,
- and industry bodies consolidate the Community's input in a transparent and efficient way.

Open Source goes beyond just code: what's next?

Participants are now able to harness the power of Goldman's internal data platform for their own data management and governance needs, not only to mutualize technology expenditure, but also to address business use cases. There are several ways to get involved:

Deploy Legend on premise. Several FINOS bank members who participated in the pilot are currently deploying Legend on premise. See legend.finos.org/docs/installation/installation-guide for instructions to deploy your own instance of Legend Studio.

Join an existing Legend modeling project. FINOS continues to host the Legend Studio shared instance. Request an account on the Legend shared instance at finos.org/legend

Propose a new, collaborative data modeling project on the FINOS Legend shared instance (FINOS members only). FINOS members can propose new Legend modeling projects at github.com/finos/community/issues/new/choose.

Contribute to Legend. Visit legend.finos.org/docs/contribute/contribute-to-legend if you'd like to contribute to further developing the Legend-PURE language and Legend platform at github.com/finos/legend.

About Legend

For more information about Legend, including the project roadmap, visit legend.finos.org, and check out the Legend code at github.com/finos/legend. Request an account on the FINOS hosted instance of legend at finos.org/legend. Subscribe for Legend updates by emailing legend+subscribe@finos.org.

About FINOS

FINOS (The Fintech Open Source Foundation) is a nonprofit whose mission is to foster adoption of open source, open standards and collaborative software development practices in financial services. It is the center for open source developers and the financial services industry to build new technology projects that have a lasting impact on business operations. As a regulatory compliant platform, the foundation enables developers from these competing organizations to collaborate on projects with a strong propensity for mutualization. It has enabled codebase contributions from both the buy- and sell-side firms and counts 38 major financial institutions, fintechs and technology consultancies as part of its membership. FINOS is also part of the Linux Foundation, the largest shared technology organization in the world.

About Goldman Sachs

The Goldman Sachs Group, Inc. is a leading global investment banking, securities and investment management firm that provides a wide range of financial services to a substantial and diversified client base that includes corporations, financial institutions, governments and individuals. Founded in 1869, the firm is headquartered in New York and maintains offices in all major financial centers around the world.

finos.org
info@finos.org
1 650 665 9773

