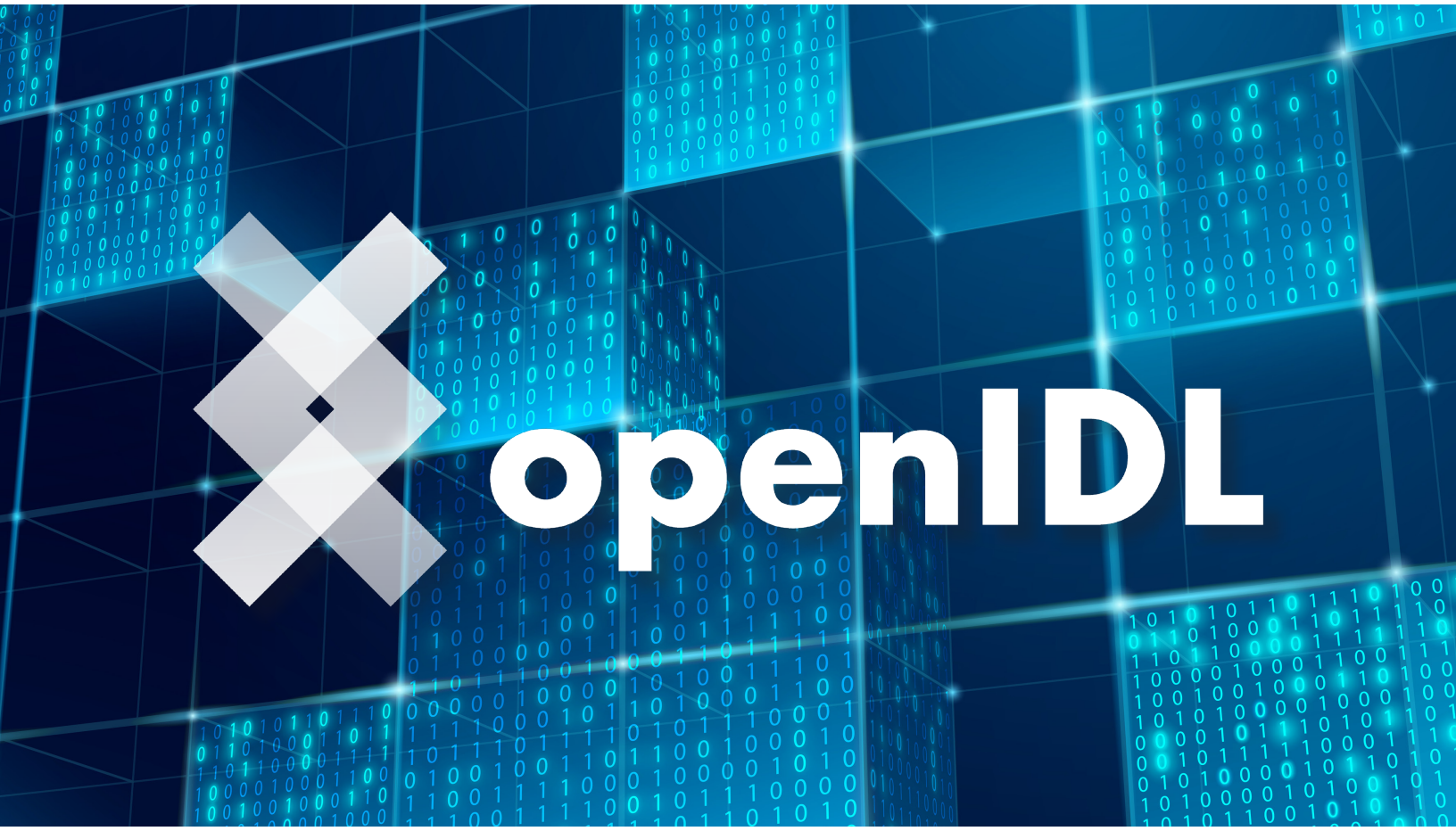
AAIS



openIDL Proof of Concept Report

Distributed Ledger Technology Passes Test En Route to Broader Industry Application

August 2021

AAIS

openIDL Proof of Concept Report

openIDL

Table of Contents

AAIS

Piloting openIDL for Insurance Regulatory Reporting

openIDL, the Open Insurance Data Link, was created by AAIS in 2018 in collaboration with regulators, insurance carriers and other insurance industry stakeholders to improve regulatory reporting and provide the foundation for an external data strategy for the insurance ecosystem. In 2021, openIDL was adopted as a Linux Foundation project to broaden its reach and impact.

Overview

In March 2020, insurance regulators and major insurers came together to answer the question: Could openIDL's blockchain/ distributed ledger technology handle current insurance regulatory reporting and improve on the process for all parties involved?

This initial Proof of Concept for openIDL centered on a COVID-19 Business Interruption Data Call issued by the National Association of Insurance Commissioners (NAIC).

The endeavor was designed to demonstrate whether openIDL could enable:

1. a data call to be executed with greater efficiency and more timely information
2. private data to be leveraged by regulators, while maintaining the privacy and security of that same data for carriers
3. greater transparency and data integrity throughout the data call process
4. deeper, more meaningful insights from data
5. openIDL's ability to correlate data from third party sources to carrier-provided data to gain new insights and inform decision making.

Results of Proof of Concept were positive, providing the evidence and confidence stakeholders needed in order to consider openIDL as integral to the improvement of regulatory reporting and secure data management across the insurance ecosystem.

AAIS

Section 1: Piloting openIDL for Insurance Regulatory Reporting

Putting openIDL to the Test

Guided by insurance stakeholders and Design Thinking sessions hosted by AAIS, initial efforts surrounding openIDL focused on building a technology platform and organizational structure that would be trusted and valued by all participants in regulatory reporting and adaptable for application development across the P&C industry. openIDL developers established an infrastructure that married the inherent benefits and security of blockchain and distributed ledger technology with the unique applications of a harmonized data store. By modernizing regulatory reporting and reframing statistical reporting with a broader, external networked data strategy, the openIDL platform could be transformational, changing the way carriers and regulators access and use data.

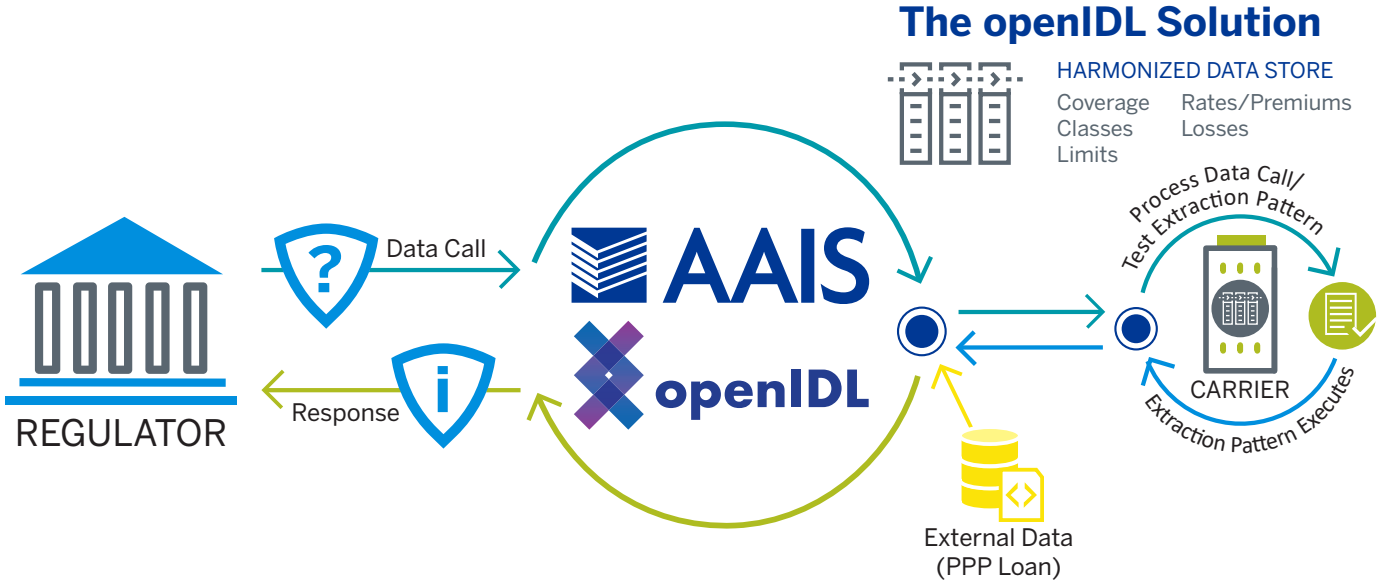
It was time to test the viability of the openIDL platform in a real-world scenario. The need for regulators to monitor the impact of the COVID-19 pandemic on the business interruption market provided the opportunity for a qualified Proof of Concept with a practical application of openIDL on a concrete data call.

Insurance Data Calls The openIDL Way

Two large insurance carriers agreed to participate in the openIDL Proof of Concept along with regulators from nine states: North Dakota, Connecticut, Mississippi, California, Texas, New Jersey, Maryland, Virginia and Alabama, as well as the NAIC. Each party to the data call was assigned a 'node' – a private participating system on the network. 'Insurer' nodes were established for carriers as the data owners, while 'Analytics' nodes were configured for the recipients of the data and information, namely the state regulators and AAIS as the licensed statistical agent.

The openIDL nodes and network were scrutinized by the insurers for compliance with security protocols. Leveraging statistical reporting data and new confidential and identifying data elements, monthly transactional data was collected in separate nodes on behalf of participating carriers. Smart contracts were deployed to keep each carrier's data private and in each carrier's direct control. In addition, the integrity of the data was assured through the openIDL network — without raw data being released or transferred.

The openIDL Data Call Lifecycle



With openIDL, carrier data never leaves their control.

AAIS

Section 1: Piloting openIDL for Insurance Regulatory Reporting

Results Exceed Expectations

The COVID-19 BI Data Call leveraged an additional data source, the Federal Government's Payroll Protection Program (PPP) loan database, during the data extraction process. Business addresses of loan recipients were correlated with policyholder addresses, supplementing policy records with loan details without exposing raw data. As a result, richer summary output information was delivered to the openIDL analytics node, allowing regulators and insurers to unlock greater insights than they could with a traditional data call – all while maintaining the privacy of carrier data – and no private or identifying data exchanged.

openIDL ALLOWED REGULATORS AND INSURERS TO UNLOCK GREATER INSIGHTS THAN THEY WOULD WITH A TRADITIONAL DATA CALL, WHILE STILL MAINTAINING THE PRIVACY OF THE DATA.

Numerous Advantages

Results of this pilot openIDL data call pointed to several advantages over traditional methods of regulatory compliance:

- Insurers could deliver far more timely data — far more efficiently. Rather than quarterly data, regulators would be able to receive monthly data via openIDL. Once the carrier node is populated, a data call that used to take as long as six months could be completed in just a few days.
- Privacy was maintained for the data owner; address and policy/claim associations remained confidential.
- The ability to associate insurer provided data with other sources, in this case PPP loan data, provided previously unattainable insights. The PPP loan information helped illuminate other insights of practical use to both regulators, insurers, and other stakeholders.
- Regulators can be supremely confident of the integrity of the information they receive knowing that the output is derived from original data sources and aligned with the transparent logic and query of their data call.

Data Call Work Flow

AAIS

THE ABILITY TO ASSOCIATE INSURER-PROVIDED DATA WITH OTHER SOURCES, SUCH AS PPP LOAN DATA, PROVIDE PREVIOUSLY UNATTAINABLE INSIGHTS.

Prerequisites for Success

Some key observations came to light from this exercise.

- Collaboration between insurers, AAIS and regulators is critical for all parties to capture the benefits of openIDL for regulatory reporting.
- A uniform open-source data standard must be applied to data prior to loading into an insurer's node in order to achieve quality results. While initially mapping data to a consistent standard requires an upfront investment of time, once established, data can be refreshed to the node with much greater efficiency for broader reporting utility going forward. Establishing a consistent data standard can also advance insurers' efforts to better manage, understand and harness data in all areas of their own operations.
- Each carrier's independent security assessment and verification is necessary so they can be confident that their node can be deployed without compromising the secure environment.

Next Steps

With openIDL now a Linux Foundation Project, AAIS and the Linux Foundation Regulatory Reporting Steering Committee will continue to work with regulators, insurers and other stakeholders to expand and evolve openIDL as the next generation standard for statutory and periodic reporting, and perhaps the foundation for broader applications across the insurance ecosystem.

Participation by stakeholders across the industry is welcomed!

The openIDL Regulatory Reporting Steering Committee (RRSC) continues to work to advance regulatory reporting with openIDL. If you would like to learn more or participate in future Proof of Concepts, please contact Robin Westcott at robinw@aaisonline.com or Lori Dreaver Munn at lorim@aaisonline.com.

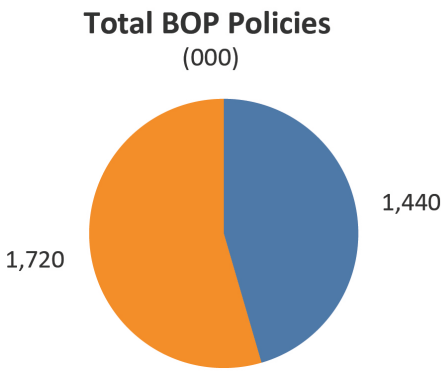
Insurer nodes enable carriers to correlate information from data fields, such as address, while keeping the data itself private and inside their companies. This game-changing approach to information-sharing allows insurers to use and associate their data with additional sources to better understand and manage their portfolios, and to provide regulators with far more meaningful and relevant insights...all in far less time.

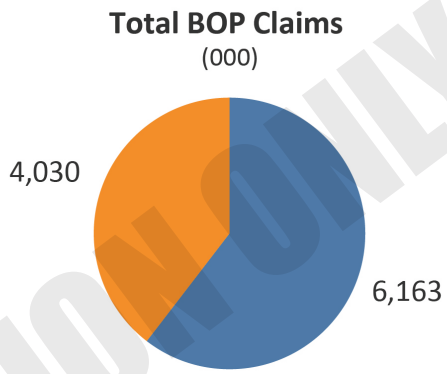
AAIS

Summary Output

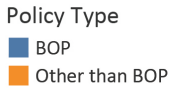
This POC output report* is intended to demonstrate that it is possible to use insurer data through openIDL to leverage other data sources to create insights previously unattainable. In this case, the third party data set used was Federal Payroll Protection Plan (PPP) loan data. The match rate between the insurer address level data leveraged for this data call to the addresses of PPP loan recipients was 93%.

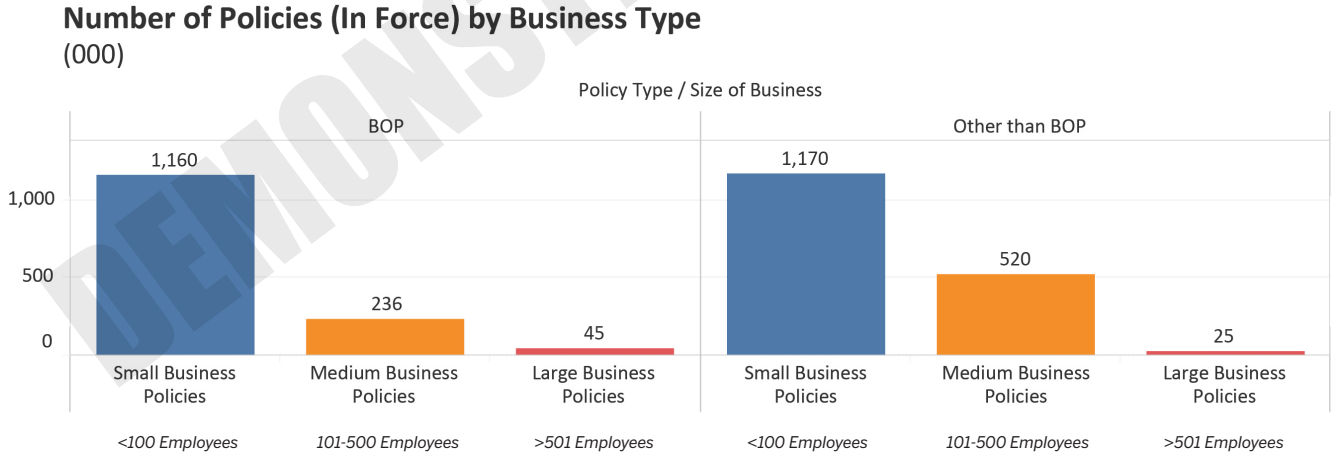
Information on policies and policy type was quickly and readily available to respond to the data call, while remaining private and in the carrier's control.





Policy Type	Total Written Premium (000)	BI Premium Written (000)
BOP	$2,216,990	$110,850
Other than BOP	$3,513,830	$175,690
Grand Total	$5,730,820	$286,540



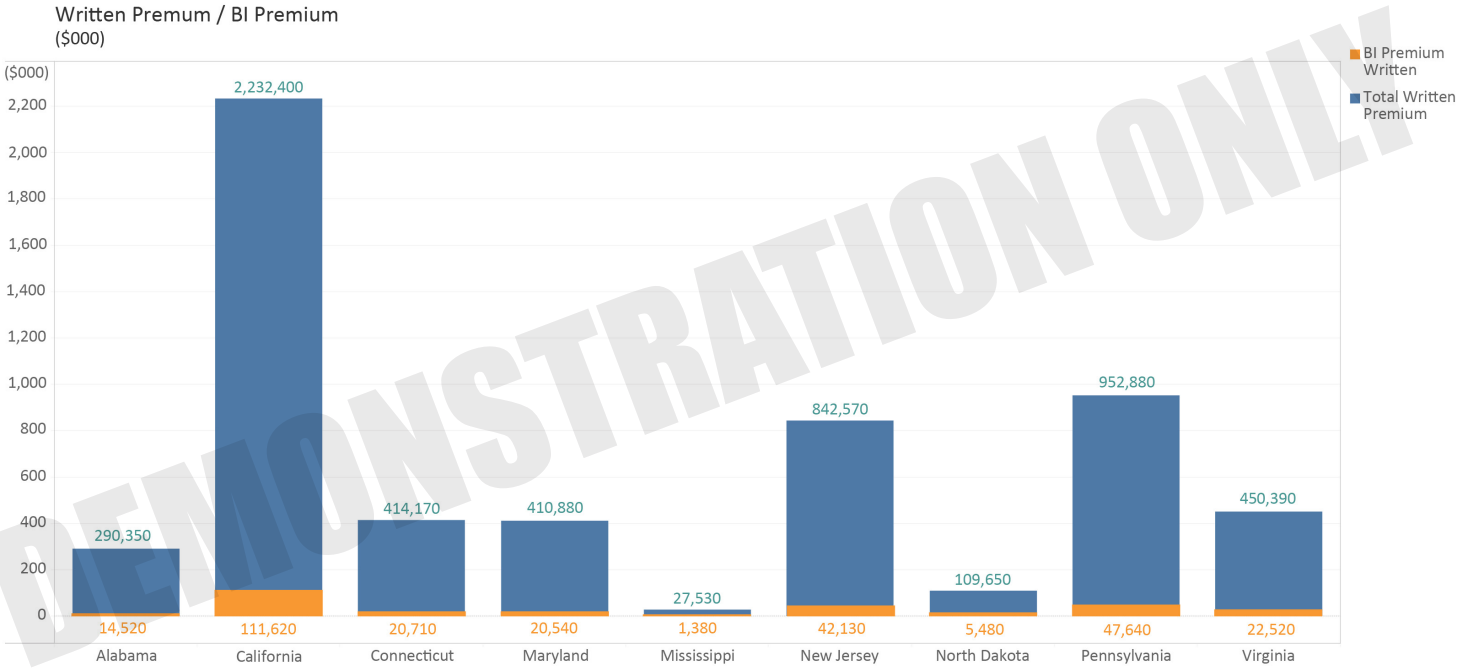


While the data used from two large carriers represented accurate data sets, AAIS added a proxy data set in order to obscure the data reported from being associated with any individual carrier. Because the results are impacted by the proxy data, the output here is for demonstration purposes only.

AAIS

Section 1: Piloting openIDL for Insurance Regulatory Reporting

Once openIDL nodes are populated, the information needed for the data call is instantaneously available, while remaining private and in the carrier's control.

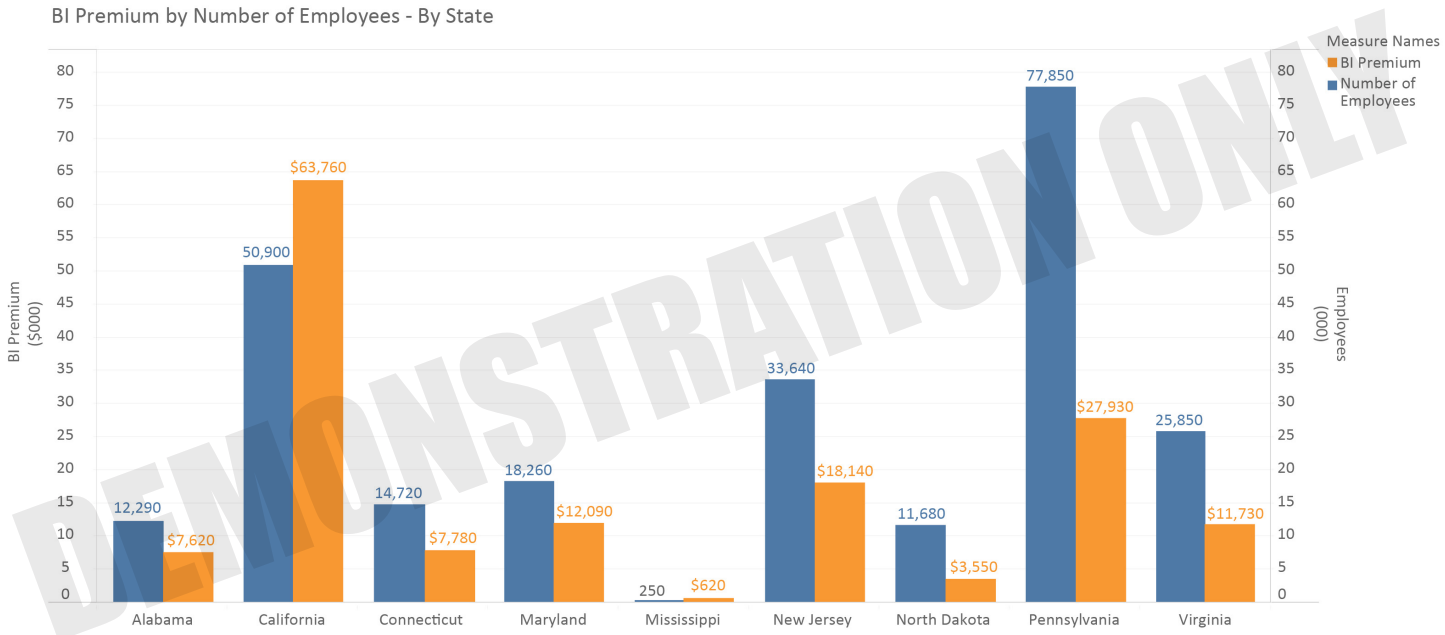


Business Interruption premium represents 5% of the total BOP premium assumption by state for participating carriers.

AAIS

Insurance Data + Third Party Data = New Insights

Correlating data in openIDL with other third party sources illuminates insights previously unattainable.



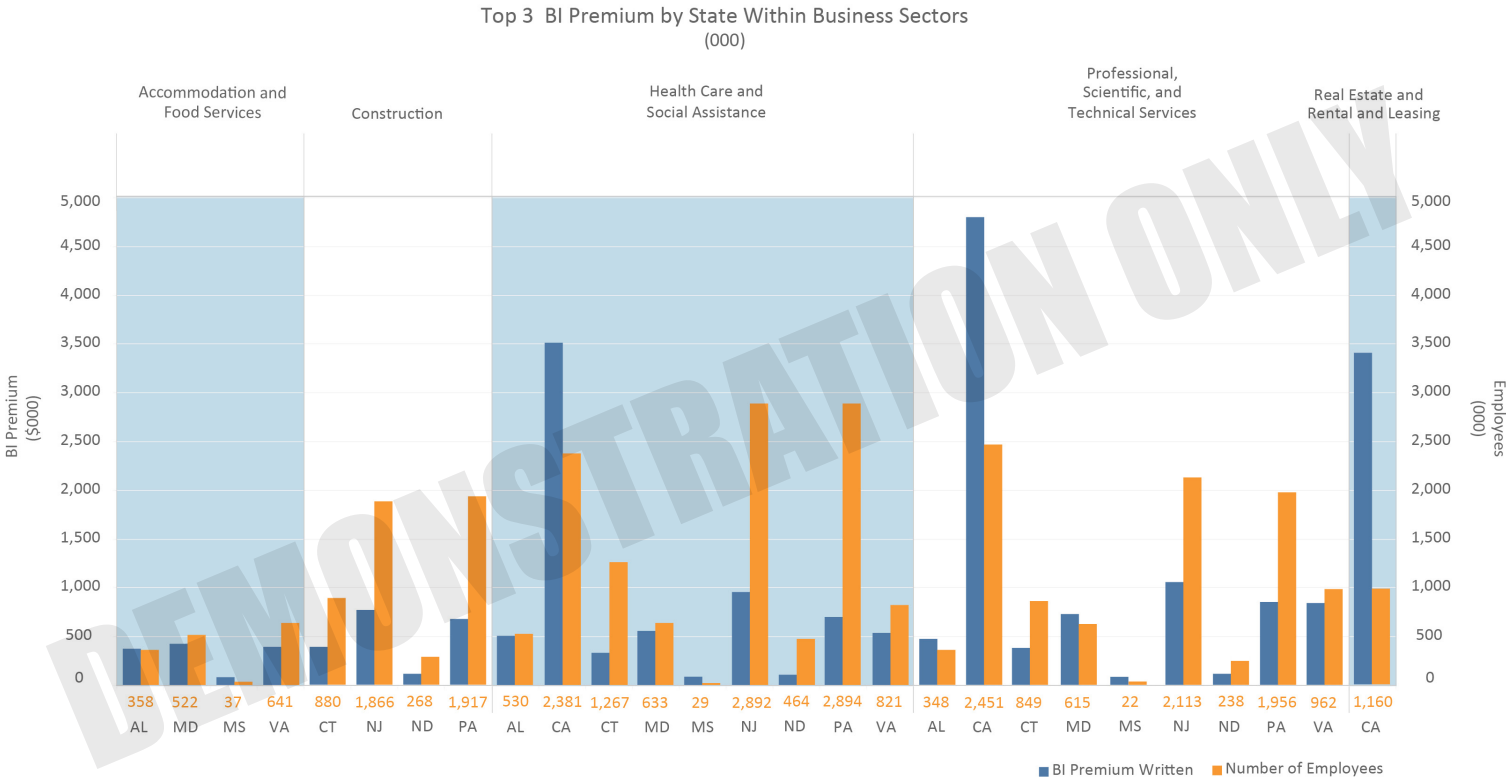
Correlating number of "jobs" reported on PPP loan applications to business/policy address revealed the number of employees per state that would be impacted by business interruption. This is a data point regulators were asking for but that, historically, without openIDL data correlations, is not readily available.

AS THIS REPORT DEMONSTRATES, THE LINK BETWEEN THE INSURED, A CLAIM AND OTHER DATA SETS CAN BE SUCCESSFULLY ENGAGED TO PROVIDE BOTH INSURERS AND REGULATORS WITH GREATER INSIGHTS, WITHOUT COMPROMISING PROPRIETARY INFORMATION OR EXPOSING CONSUMER DATA.

AAIS

Section 1: Piloting openIDL for Insurance Regulatory Reporting

Insurance Data + Third Party Data = New Insights



Correlating insurance data with PPP loan information by business/policy address revealed the number of employees that would be impacted by BI losses for each state's top three NAICS codes.

Total Business Insurance Premium Insufficient to Cover Pandemic Losses

Apparent disparity between BOP premium collected and PPP loan amounts, by state.

	AL	CA	CT	MD	MS	NJ	ND	PA	VA
BOP Premium (est) as a % of PPP Loans	0.11%	0.36%	0.05%	0.08%	0.01%	0.19%	0.04%	0.12%	0.05%

Correlating insurance data with PPP loan information revealed the amount of BOP premium collected by the insurance industry is less than 1% of the PPP loans provided to affected businesses in states listed.

The ability to correlate data from populated, private nodes with third party data sources has far-reaching implications for carriers and regulators. For example, carriers can use deeper data to better manage their portfolios and target marketing, and regulators can more effectively and efficiently attain more meaningful information for policy setting and decision-making.

Visit AAISonline for more on openIDL.

AAIS

Glossary

- Blockchain – data storage technology that leverages encrypting data between records such that it maintains the integrity of the ledger of records over time. Implied is the use of Smart Contracts and a Consensus Mechanism to assure integrity of created blocks and the ledgers across participating Peers.
- Data Call – a generic term when an insurance regulator, state department of insurance, commissioner or national body such as the National Association of Insurance Commissioners (NAIC), requests specific data of the industry. Data calls happen on an annual or scheduled basis, may be event-driven around catastrophes or context, or may be completely new, attempting to understand new or changing risks.
- Distributed Ledger Technology (DLT) – also known as blockchain, the technology that involves using ledgers with cryptographically linked records to be shared and maintained in sync across members of a network – distributed – assuring each participants of the integrity of the network, operations and data.
- Extraction Pattern – specific to openIDL, an Extraction Pattern is a set of code that is maintained through the ledger for integrity and executed on an Insurer Node when they consent to participate in a Data Call.
- ElasticSearch – a readily available and commonly used "fuzzy logic" application was leveraged for the following logical process to find common addresses between the insurer-provided Policy Address and Borrower Addresses in the Federal PPP Loan database.
- Harmonized Data Store – specific to openIDL, the Harmonized Data Store is the established database used by the node to store data that is trusted for use by the network, with consent of the owner. In this POC, it is a MongoDB datastore that is populated with data according to the format and quality rules established by AAIS and state regulators.
- "Insurer" or "Carrier" Node – a Node with a particular role as a data owner and configuration for a participating insurer, deployed by an insurer known to AAIS for statistical reporting services and participation in the POC.
- Node – a participating system on the network.
- Analytics Node – a node with a particular role and configuration as the information recipient (data destination), with multiple private channels, one with each reporting insurer.
- Insurer Node – a note to receive the results of a consented data call issued by a regulator.
- Peer – often used interchangeably with "Node," Peer is more technically identifying a participant on a blockchain network as a peer to one more node, implying role or "channel" specificity.
- Private Data Collection – a collection of data in a private state database on the peers of authorized organizations, which can be accessed from chaincode on these authorized peers. The actual private data can be sent peer-to-peer via gossip protocol to only the organization(s) authorized to see it.
- Smart Contract – a term specific to blockchain technology, a Smart Contract is the code, often called "chaincode" that is managed by the network and permitted to write new blocks to the ledger when the logic executes successfully.

AAIS

openIDL Proof of Concept Report

Section 2 - The Five Step Technical Approach

openIDL

AAIS

Section 2 - The Five Step Technical Approach

Step 1: Preparing the Network

The NAIC's 2020 COVID-19 Business Interruption Data Call was the call establishing the interaction for the Proof of Concept (POC) for industry regulatory reporting. There were four separate openIDL nodes supported, representing the three participating organizations: two large insurance carriers and AAIS, as the licensed advisory organization.

The network was prepared with the necessary applications and nodes representing each organization and their roles. Preparing Insurer and Multi-Tenant nodes to ingest data and Regulators to access the web interface to begin preparing data calls.

AAIS established the nodes in the IBM Cloud within AAIS's enterprise account, and provided complete access-control and transparency to each carrier for their respective nodes.

Step 2: Defining, Preparing & Ingesting Data

Data Elements from the NAIC Data Call

This POC used the initial elements recognized in the NAIC COVID-19 Property & Casualty Data Call. This data call was issued broadly to commercial insurers with small business owners' policies, or "BOP", premium and claim information reported separately from other commercial programs.

The data model was developed for populating the Harmonized Data Stores. The data model included data elements from the NAIC Covid Data Call, statistical data reporting as well as "Confidential Key Data".

The following fields (summarized) were requested by the NAIC and duplicated in the POC:

For Policy/Premium and Exposure:

- Total Policy Premium – Whole Policy for BOP and Commercial programs
- Business Interruption (BI) Premium (if separable)
- Number of Policies in Force as of Dec. 31, 2019
- Number of Policies with Exclusion for Pandemic (et al)
- Number of Policies with Physical Damage Requirement

For Claims and Related Losses:

- Claims and Losses related to COVID-19 – directly related to the disease, civil authority actions or where state or federal legislation requires the insurer to pay.
- Loss – indemnity payments on claims, excluding adjustment expense and subrogation and salvage.
- Payment – loss payments, not including adjustment expenses.
- Case Incurred Loss – case reserves plus claim payments made to date.
- Paid Losses – total paid losses on closed claims.

AAIS

Part 2: The Five Step Technical Approach

Available Data: Previously-reported Statistical Data

Participating insurers already perform statutory reporting through AAIS as their statistical agent and provide Policy (exposure/premium) and Claim (loss) information in an unrelated fashion (so claims and policies cannot be linked), and at a ZIP code level for the risk geographically. The following data points were already understood and readily available from participating carriers through existing processes:

Policy Record Data Elements	Claim (Loss) Record Data Elements
Line of Insurance	Line of Insurance
Accounting Date	Accounting Date
Company Code	Company Code
State Code	State Code
Transaction Code	Transaction Code
Premium Amount	Loss Amount
Exposure	Claim Count
Annual Statement Line of Business	Annual Statement Line of Business
Policy Form	Policy Form
Coverage Code	Coverage Code
Class Code	Cause of Loss
ZIP Code	Accident Date
Policy Number	ZIP Code
	Claim Number
	Claim Identifier

Establishing the Confidential Keys: Policy Address and Policy ID on Claims

Neither previously supplied data for statutory reporting nor data reported for the NAIC COVID-19 Data Call were adequate to provide any more value than a traditional data call. In order to improve insights, this POC would need to identify, permit and privately access other key data fields.

Because we were requesting access to data typically kept private by enterprises, the bulk of the time of the POC was spent working through data definition questions and researching points of data to harmonize, while adhering to enterprise security protocols permitting the use of production data and technical participation in the POC.

Many identifying data points were researched for potential use. Most were dismissed largely due to the inability to access them in a consistent and timely fashion -- if they were captured by the insurer at all. These included identifiers like Federal Tax ID (FEIN), Standard Industrial Classification (SIC) codes, detailed risk location address and Number of Employees.

As a result of this work, Insurers in the POC were asked to supply a few key fields to enhance their previously-reported Policy/Premium and Claims/Loss experience data. Specifically to:

1. **Add "Policy ID" to Claims records – linking specific claims to policies**
2. **Add "Policy Address" to Policy records – the address of the legal entity or person taking out the policy; not to be confused with the specific risk location (insured building address) which may be different.**

AAIS

Part 2: The Five Step Technical Approach

Leveraging Third Party Data: The Federal Payroll Protection Program (PPP) Loan Database

Proving the ability to leverage private data required relating the private key data to an additional data source to enhance the details of our policy and claim records.

The Federal Payroll Protection Program (PPP) Loan database had address level information for businesses accessing PPP loans. The publicly available Federal PPP Loan database https://data.sba.gov/dataset/ppp-foia also includes many useful details on the loans and those businesses (and owners) applying for them, and AAIS selected the following for inclusion in the POC operations:

- Borrower Address: Street Address, City, State and ZIP
- Initial Approval Amount (of PPP Loan)
- Current Approval Amount (of PPP Loan)
- Undisbursed Amount (of PPP Loan)
- Servicing Lender Location ID (unique identifier to lender)
- Servicing Lender Name
- Rural/Urban Indicator
- Hubzone Indicator
- Business Age Description
- Congressional District
- Jobs Reported – the number of jobs the PPP Loan is purported to impact
- NAICS Code – the 6 digit industrial classification code
- Originating Lender State
- Non-Profit Business – "Yes" if applying business is a non-profit Org

ASSOCIATING DATA ON PPP LOANS TAKEN BY BUSINESSES AND INDIVIDUALS FOR THEIR COMPANIES, THROUGH ADDRESS, TO BUSINESSES AND INDIVIDUALS WHO HAVE PREVIOUSLY TAKEN OUT INSURANCE POLICIES, ENABLED GREATER INSIGHTS TO BE REGULATORS, WHILE KEEPING PRIVATE DATA PRIVATE AND PARTICIPATING INSURERS ANONYMOUS.

Step 3: Loading Harmonized Data Stores

The three participating organizations, Insurer 1, Insurer 2 and AAIS, on behalf of many insurers, delivered and ingested data to their Harmonized Data Stores slightly differently, enabling efficient paths for each enterprise to participate. Each organization was able to establish the key Policy and Claim data points in their Harmonized Data Stores, keeping the data private, anonymous, and assured of integrity in anticipation of the regulatory data call.

- Insurer 1: Re-submitted (updated) previously supplied statistical data, with the additional columns for Policy Number on Claim records, and Policy Address on Policy records.
- Insurer 2: Supplied key-pair fields only, Policy-Policy Address and Claim-Policy Number, so previously supplied data could be leveraged for the data call.
- Multi-Tenant (AAIS) Node: In order to supply the POC additional data to assure participating insurers of anonymity, AAIS leveraged previously reported statistical data from dozens of companies, and notionally (actuarial term for "artificially," "falsely" or otherwise made-up data) and randomly associated PPP Loans to those policies.

Step 4: Issuing the Data Call & Extraction Pattern

Regulators worked with AAIS to draft an updated COVID-19 Business Interruption (C19 BI) data call and issue it for the participating states through the openIDL interface.

AAIS updated the Data Call with the Extraction Pattern, which is saved to the ledger, assuring nothing but that code is executed, and the data call was ready for review and "consent" by an insurer. Specific to openIDL, an Extraction Pattern is a set of code that is maintained through the ledger for integrity and executed on an Insurer Node when they consent to participate in a data call.

AAIS

Part 2: The Five Step Technical Approach

The Extraction Pattern was three part:

1. the query of insurer data, as a "Map/Reduce" function, and
2. a small, commonly-used, transparent ElasticSearch[1], delivered to the node to perform a "fuzzy logic" match on the business/policy address. This permitted the PPP data points to enhance the policy data in the network without any address or private information leaving the insurers control. This resulted in a high quality and high frequency of matching addresses in testing.
3. the enrichment of data occurs such that the new private details of the policyholder from the loan information can be kept private from the insurer themselves. Sensitive details may not be disclosed at a policy level - but could be reported in aggregate or across a relevant dimension for the benefit of the insurer.

Step 5: Delivery to the Analytics Node

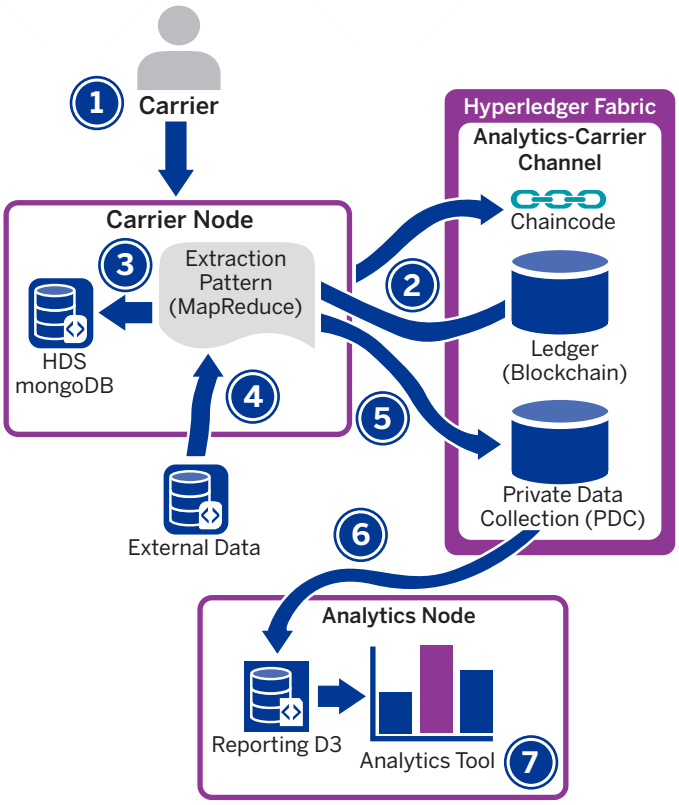
Once the The POC Data Call was "Issued", and the Extraction Pattern executed, the data was returned and the identity of the resulting data set was hashed to the ledger and placed in the respective insurers' Private Data Collection (PDC) of their Hyperledger Fabric peer. The data was replicated from the insurer PDC to the PDC on the Analytics Node through a Gossip protocol replicating the data securely, reliably and accurately, by validated through the ledger.

Extraction Patterns, Chaincode and the PDC Deliver Data to Analytics Node

Data gathered by the Extraction Patterns on each consenting node is securely replicated to the Analytics Node through Hyperledger Fabric's Private Data Collections and the GOSSIP protocol.

Extraction Pattern

1. Carrier Consents to Data Call
2. Extraction Pattern Retrieved from Ledger.
3. Extraction Pattern Run Against Harmonized Data Store
4. Additional Decoration of Data with External Data can happen here...like PPP
5. Result of Extraction Pattern Placed into Private Data Store
   - Hyperledger Fabric replicates for Analytics Node
6. Result Loaded into Reporting DB on Analytics Node.
7. Analytics Tool Generates Report.



[1]ElasticSearch, a readily available and commonly used "fuzzy logic" application was leveraged for the following logical process to find common addresses between the insurer-provided Policy Address and Borrower Addresses in the Federal PPP Loan database.

AAIS

Part 2: The Five Step Technical Approach

Data from participating Insurer and the Multi-Tenant Nodes was aggregated to the local Harmonized Data Store on the Analytics Node.

The resulting data set had the following elements:

- State Name
- Line of Insurance
- Annual Statement Line of Business
- Policy Form
- Major Peril
- Number of Employees (not populated)
- Physical Damage Requirement (not populated)
- Transaction Code
- Transaction Name
- Coverage Code
- Cause of Loss Code
- Cause of Loss
- Claim Status
- Accounting Date
- Loss Accident Date
- Business Interruption Flag (BI Coverage, assumed "Y" for BOP, or "N" for other)
- Physical Damage Requirement Flag (Assumed "Yes")
- Viral Exclusion Flag (Assumed "Yes")
- SIC Code (not populated)
- PPP Indicator (Y if a matching PPP Loan is found; N otherwise)
- NAICS Code (from matching PPP Loan)
- Loss Amount
- Written Premium
- Number of Policies (in "chunk")
- Number of Claims (in "chunk")
- Chunk ID – identity of the hash attesting the integrity of the row

AAIS then prepared the aggregated data set into a Tableau Workbook for reporting and visualization. See *Summary Output,* for demonstration purposes only, starting on page 7.

Visit AAISonline for more on openIDL.