

# 2022-7-15 Meeting Agenda

## Date

15 Jul 2022

## Antitrust Policy Notice

Linux Foundation meetings involve participation by industry competitors, and it is the intention of the Linux Foundation to conduct all of its activities in accordance with applicable antitrust and competition laws. It is therefore extremely important that attendees adhere to meeting agendas, and be aware of, and not participate in, any activities that are prohibited under applicable US state, federal or foreign antitrust and competition laws.

Examples of types of actions that are prohibited at Linux Foundation meetings and in connection with Linux Foundation activities are described in the Linux Foundation Antitrust Policy available at <http://www.linuxfoundation.org/antitrust-policy>. If you have questions about these matters, please contact your company counsel, or if you are a member of the Linux Foundation, feel free to contact Andrew Updegrave of the firm of Gesmer Updegrave LLP, which provides legal counsel to the Linux Foundation.



This is a weekly series for The Regulatory Reporting Data Model Working Group. The RRDMWG is a collaborative group of insurers, regulators and other insurance industry innovators dedicated to the development of data models that will support regulatory reporting through an openIDL node. The data models to be developed will reflect a greater synchronization of data for insurer statistical and financial data and a consistent methodology that insurers and regulators can leverage to modernize the data reporting environment. The models developed will be reported to the Regulatory Reporting Steering Committee for approval for publication as an open-source data model.

openIDL Community is inviting you to a scheduled Zoom meeting.

Join Zoom Meeting

<https://zoom.us/j/98908804279?pwd=Q1FGcFhUQk5RMEpkaVIFTWtXb09jQT09>

Meeting ID: 989 0880 4279

Passcode: 740215



GMT20220715-1...1920x1080.mp4

One tap mobile

+16699006833,,98908804279# US (San Jose)

+12532158782,,98908804279# US (Tacoma)

Dial by your location

+1 669 900 6833 US (San Jose)

+1 253 215 8782 US (Tacoma)

+1 346 248 7799 US (Houston)

+1 929 205 6099 US (New York)

+1 301 715 8592 US (Washington DC)

+1 312 626 6799 US (Chicago)

888 788 0099 US Toll-free

877 853 5247 US Toll-free

Meeting ID: 989 0880 4279

Find your local number: <https://zoom.us/j/98908804279>

## Attendees

- [peter antley](#)
- [Eric Lowe](#)
- [Dale Harris](#)
- Mike Nurse
- Jenny Tornquist
- Wanchin Chou
- Andy Case
- Reggie Scarpa
- [Ash Naik](#)
- [Ken Sayers](#)
- Bourjali Hi
- [Kristin McDonald](#)
- [Susan Chudwick](#)
- [James Madison](#)
- [Jefferson Braswell](#)
- [Nathan Southern](#)
- [George Bradner](#)

## Goals

## Meeting Minutes

I. Opening

A. Greetings to attendees

B. Mr. Antley: LF AntiTrust statement.

II. Agenda - Data Modeling and Internal Data Work - Mr. Antley

A. Overview/Recap

1. Discussion of data modeling and what idl project is doing with data. Concerned that discussions may be becoming overly technical. Wants to dial the tech speak down a bit
2. Pulled up graphic on Earned Premium - 3 columns: Input / Business Layer /Output
3. Noted that last week, Mr. Williams led a discussion on car years.
4. Also noted that he led the discussion on business layer - wants to revisit and continue this.

#### B. Layers of Current Model (Input/Business Layer/Output) + Transformations

1. Re: Input, Stat plan as it gives us adequate data to do regulatory reporting, and we can ingest most of needed information:
  - a. Accounting date
  - b. Duration
  - c. Coverages and exposures
2. Re: Output: some of the terms for which people are asking (presented examples here): e.g., Earned premiums, car years, incurred loss, earned /written exposures and premiums, incurred claims/losses, calculated ratios, average expenditures, average premiums, underwriting expenses.
3. Current transformations
  - a. Mr. Antley: the way they solve (@AAIS, in his warehouse) these questions today: they take the input records, process these to be centered around the idea of a policy, and then calculate: earned premium, paid loss, incurred loss, outstanding losses, earned exposures, paid claim counts, outstanding claim counts. They calculate all of this on quarterly basis.
  - b. Sits and persists in data store on level on which they are making reports, they have these terms, grouped by policy, calculated by quarter. If x person asks "what is my earned premium for the year, for these kinds of coverages," he does a select query that does filters and aggregates and produces year output. Data on a policy and quarterly basis with the premium after it has been earned.
4. HDS - In AWG: Discussion - how do we set up HDS so it is accessible to regulators, sans putting undue burden on rest of system?
5. Pointed to base layer, from which we answering questions. How do we make it accessible to the persons making queries? Quarterly basis reports right now - fairly easy and straightforward
6. Mr. Antley solicited thoughts/suggestions from group about what the data layer will look like
  - a. Mr. Sayers: when looking at data layer, we must ask visibility to whom? Carriers e.g., want to see stat plans eventually with more data. Can we put this in HDS per Mr. Braswell's suggestion? Data layer could be an ephemeral pass-through that is part of the extraction pattern, or a standing entity where ETL transforms stat plan. Key questions: what are we asking the carriers to prove, and do they care about any format other than the stat plans?
  - b. Mr. Antley: who are the voters in this RRDWGW? We have carriers, and we have regulators. Another critical question: if we just look at the easiest way to load this, and say, the carriers can just be done with loading the stat records, at this point who is tasked w/turning the stat records into a query that makes sense for the regulators? Without relying on regulators to expend the energy/effort into turning records into something useful?
  - c. Mr. Sayers: There is discovery here. We start from the stat plan - may want to make it less codified to make it more workable, but it can be a perfectly viable raw material. We can throw some reports at it however and say how difficult is it to get from the stat plan data to this report? Points to ultimate question: do we need an intermediate format?
  - d. Mr. Madison: In this meeting, the intermediate business layer is as flexible and organic and transformative as we need it to be. Much more critical to define input and output layers. Stat plan = Day 1 (broad consensus), and then if we add numerous additional elements to this input layer, we can also make considerable progress. There also seems to be a series of additional elements that we want in the output layer that are not in the input - Simplest model complete conformity between input/output elements, but we're talking about something more complex and sophisticated than this. If something is a raw element, it has to go on input side. If it's stat plan plus, still has to go on left hand side - this is a column that will grow in complexity as the years pass. Output elements should be those that are fundamentally derivations from layer 1. Business layer is an unknown for now.
  - e. Mr. Braswell to Mr. Antley: this group should remain focused on data modeling, and not the machinations of transforming data through various stages. This falls under the aegis of the Architecture Working Group.
  - f. Mr. Madison agreed with this point but distinguished the business machinations - re: what we have in input vs. output. (Mr. Braswell agreed with this point fully). This meeting is not trying to solve what is in business layer.
  - g. Mr. Harris: re: using existing stat plan as a means of building the plumbing. We spent months putting together auto data model. Much time being spent on trying to engineer existing stat plan. he is more concerned that plumbing itself is working from an architectural perspective. i.e., request comes in, permission gets approved, the extraction occurs and at the end of the day we have a report. In his mind: just the plumbing. Called for us to go back to the data model we have and into which we invested a great deal of time. Argued this is most pertinent.
  - h. Mr. Braswell: to Mr. Harris's point this is much more of an architectural concern, but idea of using stat plan to move things forward and make sure it all works is an excellent one.

- i. Mr. Antley: having a great challenge making desired outputs from inputs. Mr. Harris: it works in AAIS with existing stat plan. Why can't it work in idl?
  - j. Mr. Braswell: this is expressly an architectural question. AAIS work involves many workflow steps - this needs to be respread out now over source and destination.
  - k. Mr. Sayers: extraction pattern has been broken up into multiple levels to get it to work in AAIS. Putting all of those layers into one extraction pattern is extraordinarily complicated.
  - l. Mr. Madison: there is already considerable work (modeling) going on in transformation from left (input) to right (output) - two things happening simultaneously. if we knead them together it gets muddled. One of questions: what are rules happening in middle layer. Another what is technical design of this that allows efficient adherence to those rules? Clarity of the rules is a critical question and needs elucidated. It is possible to overengineer any model: no such thing as a right level normalization. 0-6 with one layer in the middle with 8 levels. We need to clarify what we are trying to achieve. It is necessary to have the logic in the middle layer, and data structure and technology in the middle. Middle layer contains modeling but also business rules. We can't mash these together.
  - m. Mr. Braswell: with this in mind, we have to identify elements on right side that are not carbon copies, and that require additional legwork for transformation
  - n. Mr. Madison: this is question. Not a single straight move of data left to right in every case, for instance, consider earning premium, and developing losses. These are more byzantine. We should identify these elements.
  - o. Mr. Antley: is there any single straight move? Others: straight moves would include summing/aggregates. So yes there are.
  - p. Mr. Madison: maybe from a business perspective, but less of a straight move from Mr. Antley's perspective however when we are summing across 13 different stat plans.
  - q. Mr. Lowe: you should be summing in just one table for purposes of producing output for private passenger auto vs. homeowners, which are coming in on 13 different feeds
  - r. Mr. Madison: This is where ETL starts. Summing up by Mr. Antley across stat plans, and it is necessary to pull from different tables.
  - s. Mr. Sayers and Mr. Lowe: written exposures and earned exposures are a straight move across, but written/earned premium are calculated
  - t. Mr. Harris: but by the time the stat reports are created (2 years late) written premium = earned premium so no calculation necessary.
  - u. Group: with greater frequency of stat collections mean calculations are necessary
  - v. Mr. Madison: Revisiting the middle column/business layer. We have to build a better normalization model which would reduce the time commitment. Then we need to have the ownership discussion - i.e., where does the data go? On carrier nodes or elsewhere (e.g. analytics node)? This is a different discussion? But shape of data we can't know until we look at right hand side and ask what it takes to get there. All starts with the rules for anything on the right that isn't a straight move. Then we have a target and a sense of just how normalized it needs to be.
7. Mr. Harris - shared his screen re: answering a question previously introduced by Mr. Madison
- a. Looking at calls Travelers does - also including stat reporting - presented aggregations travelers needs to put together for data calls. These are the answers re: what info we need to aggregate across all lines of business to get to middle block
  - b. This is information he got from data calls Travelers has to perform, and their stat plan. Items in yellow and brown are the ones that need to be captured. (Written premium, policy counts, policy in force, etc.)
  - c. Mr. Madison: point of clarification - written premium is available on stat plan model (yes). Stat 1.0 - no aggregation needed, just carries over. (For written premium). Earned premium not in stat model. Will we put it in stat model and tell carriers they have to generate it? (No). So earned is "derived" from 1.0 data. This means written premium would be in base model.
  - d. Mr. Madison: Installment charges from Mr. Harris's table - not in the Base 1.0 model. May be in the Base 2.0 model. Could we derive this from current stat model? No. Can't be derived at all. So it should be "Base 2.0" Not available in any of plans we have today.
  - e. Mr. Madison praised Mr. Harris's table as an outstanding artifact, and proposed that we codify it as the collective requirement set - classifying each element as base or derived will tell us exactly what we need to do in terms of getting derived rules/effectuating transformations. (Others broadly agreed).
  - f. Mr. Antley: Months covered - basic aggregation, Mr. Madison stated derived 1.0
  - g. Mr. Antley: Policy counts - sum of policy #s - derived 1.0 (Ms. Chudwick). Same for coverage counts. Mr. Madison disagreed - this went in as base 2.0.
  - h. Mr. Lowe: Policies In Force is a curious one because that is a snapshot in time. That means: as of today. (Others agreed). Mr. Harris: yes, as of a point in time. Mr. Lowe: we identify the beginning and end points and as regulators calculate the figure from the delta point. Mr. Madison: this should be Derived 2.0
  - i. Mr. Antley: Claim counts definitely a sum but should be base 1.0 because we have claim # or identifier, just a matter of counting that. Need to be counted up across carriers
  - j. LAE & ALAE (Loss Adjusted Expense, Allocated Loss Adjusted expense) - can be directly assigned to a claim. ULAE: Claims manager's salary split across all claims. Mr. Harris: within our statistical report Travelers provide ALAE as part of losses. Doesn't provide ULAE. Mr. Lowe: 'Incurred Losses' a function of ALAE & the losses as well doesn't include ULAE. Agreement from group: all of these should be in Base 1.0 model - ALAE, Incurred Losses & Paid Losses.
  - k. Mr. Lowe: Incurred usually factored w/allocated reserves - your prediction of what you are going to pay. So it should be Derived 1.0.
  - l. Mr. Lowe: Every carrier's ULAE factor is different based on whether you're overpaying or underpaying your employees
  - m. Mr. Antley: Accident Date - Base 1.0
  - n. Mr. Antley: Accounting Period: Depends on the data call. They are looking for a specified time frame (calendar year, quarter, etc.) Working off the accounting date that is given - building off of another field. Base 1.0
  - o. Mr. Antley: Calendar year grabbed off of Accounting Date as well so this is base 1.0
  - p. Mr. Antley: Policy Effective Date - Base 2.0 because we're asking for policy effective and coverage effective and expiration date.
  - q. Mr. Lowe: Policy year as well base 2.0 because we need policy effective date to figure out what it will be.
  - r. Ms. Chudwick: The only way to derive earned premium - in lieu of Derived 1.0, the only way to derive for ace to derive the earned premium is to assume mid month. So have true EP derived 2.0 would have to be updated to use the accurate policy effective date.
8. Mr. Madison: Looking ahead to where we are going
- a. Mr. Antley's report... as an example. Rather than worrying about format... We should be able to take Mr. Antley's report... One of rules of design in a situation like this is use the reports. But can we flip the report into another column and place it in Mr. Harris's spreadsheet (here Mr. Madison referenced columns I-K), in same basic format, but talking about which data elements it uses?
  - b. This would mean identifying say 12 elements from Peter's table necessary to generate the table, then doing a filter/aggregation and asking if the elements are accounted for in Mr. Harris's table and simply add the ones that are not. We then take the 12 elements, classify as Base 1.0/Base 2.0, Derived 1.0/Derived 2.0. Any that say Derived 1.0, we must have the rules (formula etc.) We can say to Mr. Antley "go trace code. Find it." Other possibility is: no, we're defining the rules right here, and then derive it from there. We must make a decision for every single one of the Derived 1.0s. Once we define those Peter can almost ignore the back end. In an instance where we have no idea how to do Derived 1.0, Mr. Antley has to go through the code and locate and identify the rule.
  - c. Mr. Lowe: this is essentially the same point David Reale made earlier in the week: how do we know we have the right calculation formula if we try to back into it? (Mr. Madison agreed) rather than saying this is what we know we want and this is how it's calculated.

- d. With this in mind, Mr. Madison proposed a migration/interpolation of Mr. Antley's report into Mr. Harris's table (the others broadly agreed), followed by a determination about which are Derived and identification of the pertinent calculations. For all of the Derived 1.0s we must have rules.
- e. Mr. Antley: amenable to this but per Mr. Madison's suggestion let's just resume that discussion offline and finish remainder of categorizations in Mr. Harris's table for now.
- 9. Discussion of Harris table/related categorizations, Part 2.
  - a. Annual Statline - Base 2.0.
  - b. Financial Line - Stat Line Extension (subline personal auto) - Line 19.2 on Annual Statement - Private Passenger Liability. Same as Annual Stat Line, redundant, so we can eliminate
  - c. Coverage - Base 1.0.
  - d. Class Code - Base 1.0
  - e. Policy Form - internal issue - type of product travelers provide. helps actuary group for rate making purposes.
  - f. Limit - Base 1.0
  - g. Deductible - Base 1.0
  - h. State - Base 1.0

Discussion items

Time	Item	Who	Notesat

Notes

Action items

☐