

Semantic Metadata **Annotation** for Network **Anomaly** Detection

draft-netana-opsawg-nmrg-network-anomaly-semantics-01

Helps to test and validate outlier detection, supports
supervised and semi-supervised machine learning development
and make anomalies for humans apprehensible

thomas.graf@swisscom.com
wanting.du@swisscom.com
alex.huang-feng@insa-lyon.fr
vincenzo.riccobene@huawei-partners.com
antonio.roberto@huawei.com

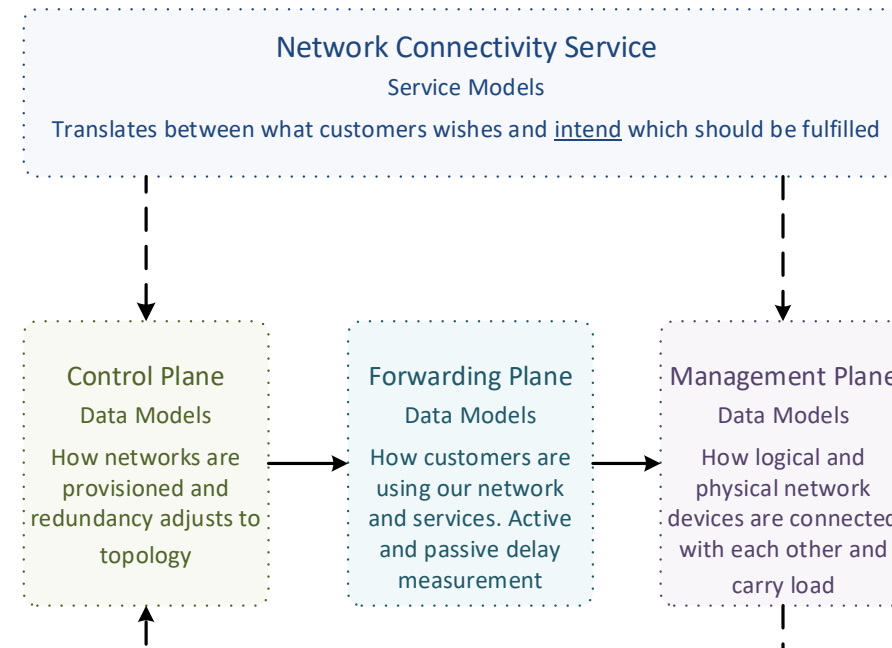
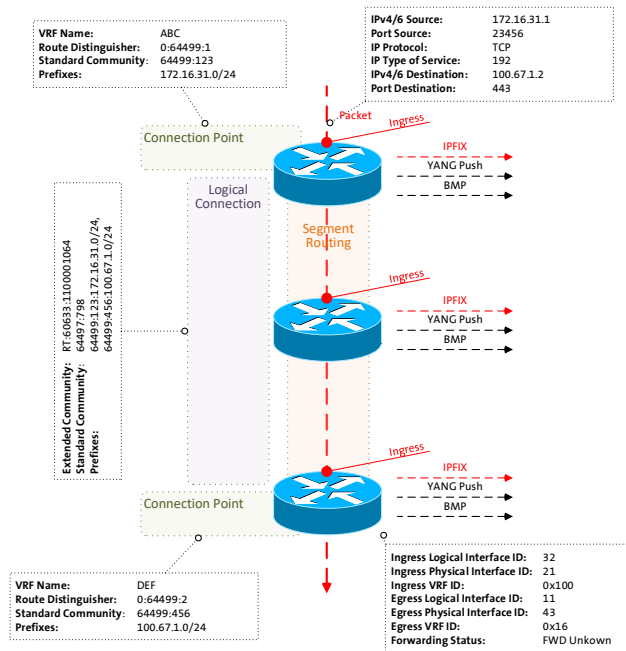
29. October 2023

Semantic Metadata Annotation for Network Anomaly Detection

What monitor and which operational metrics we collect

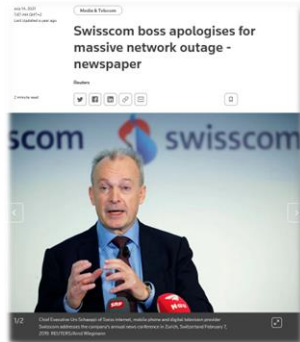
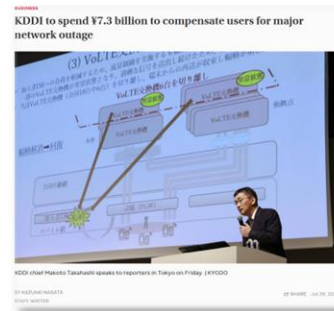
« Network operators **connect customers in** routing tables called **VPN's** »

« Network Telemetry (**RFC 9232**) describes how to collect data from **all 3 network planes** efficiently »



Semantic Metadata Annotation for Network Anomaly Detection

Recognize network incidents faster than humans can



05 FEB 2023 | 08:23 AM UTC

Italy: TIM internet services interruption reported nationwide Feb. 5

TIM internet services interruption reported in Italy Feb. 5. Likely communication disruptions.

Informational Communications/technology Transportation ITA



Facebook outage: what went wrong and why did it take so long to fix after social platform went down?

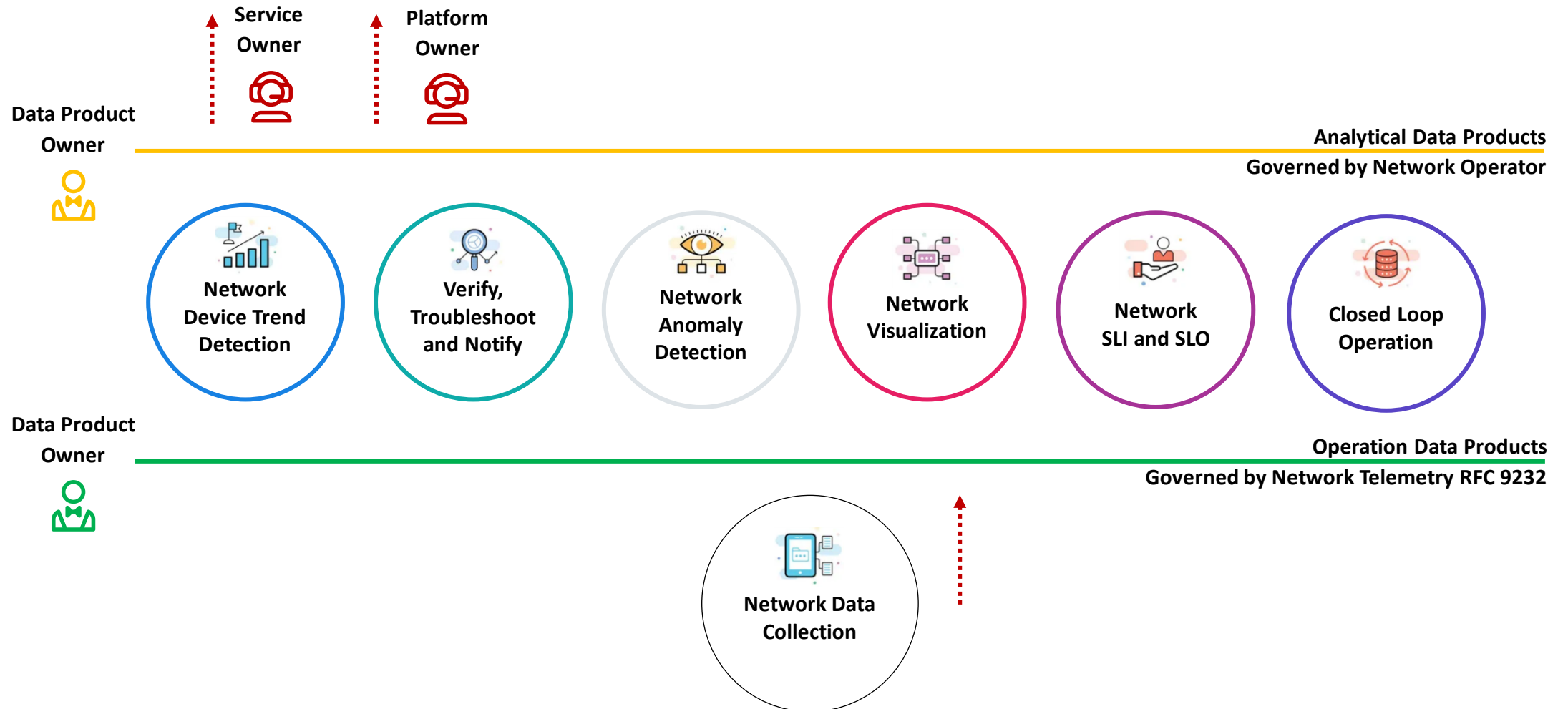
Billions of users were unable to access Facebook, Instagram and WhatsApp for hours while the social media giant scrambled to restore services.



« Customers are **always connected**, when **VPN's changing**, regardless due to operational or configurational reasons, network operators are **late to react** due to **missing visibility and automation** »

Semantic Metadata Annotation for Network Anomaly Detection

The Data Mesh Architecture enables Network Analytics use



Semantic Metadata Annotation for Network Anomaly Detection

What does Network Anomaly Detection mean?



Network Anomaly Detection

For VPNs, Network Anomaly Detection **constantly monitors and detects any network or device topology changes**, along with their associated forwarding consequences for customers as outliers. Notifications are sent to the Network Operation Center before the customer is aware of service disruptions. **It offers operational metrics for in-depth analysis**, allowing to understand on which platform the problem originates and facilitates problem resolution.



Answers

What changed and when, on which connectivity service, and how does it impact the customers?



Focuses

Provides meaningful connectivity service impact information before customer is aware of and support in root-cause analysis.



Data Mesh

Consumes operational real-time Forwarding Plane, Control Plane and Management Plane metrics and produces analytical alerts.



Direction

From connectivity service to network platform.

Semantic Metadata Annotation for Network Anomaly Detection

Our Motivation

From network incidents postmortems we network operators **learn and improve**, so does network anomaly detection and supervised and semi-supervised machine learning.

The more network incidents are observed, the more we can improve. With more incidents the **postmortem process needs be automated, let's get organized** first by defining human and machine-readable metadata semantics and annotate operational and analytical data.

Let's get further organized by exchanging standardized labeled network incident data among network operators, vendors and academia to **collaborate on academic research**.

« The community working on Network Anomaly Detection is probably the only group **wishing for more network incidents** »

Semantic Metadata Annotation for Network Anomaly Detection

What is an **outlier** and how do we **categorize** them

Global outliers: An outlier is considered "global" if its behaviour is outside the entirety of the considered data set.

Contextual outliers: An outlier is considered "contextual" if its behaviour is within a normal (expected) range, but it would not be expected based on some context. Context can be defined as a function of multiple parameters, such as time, location, etc.

Collective outliers: An outlier is considered "collective" if the behaviour of each single data point that are part of the anomaly are within expected ranges (so they are not anomalous in either a contextual or a global sense), but the group, taking all the data points together, is.

« **Collective outliers** are important since networks are connected. Through **different planes** we observe **interconnected** symptoms from various angles »

Semantic Metadata Annotation for Network Anomaly Detection

What is a **symptom** and how do we **categorize** them

Action: Which action the network node performed for a packet in the forwarding plane, a path or adjacency in the control plane or state or statistical changes in the management plane.

Reason: For each action one or more reasons describing why this action was used. From drop unreachable, administered, and corrupt in forwarding plane, to reachability withdraw and adjacency teared down in control plane, to Interface down, errors or discard in management plane.

Relation: For each reason one or more relation describes the cause why the action was chosen. From missing next-hop and link-layer information in forwarding plane, to reachability withdrawn due to peer down or path no longer redistributed.

« Symptoms are categorized in **which plane** they have been **observed**, their **action, reason and cause** »

Semantic Metadata Annotation for Network Anomaly Detection

Do you care? Questions to the audience

Network Operators: Do you agree that today actions, traffic is dropped, path is withdrawn, interface is down are always exposed through Network Telemetry, but reason and cause, dropped due to unreachable next-hop, withdrawn due to peer down, interface down due to missing signal, are rarely and would be most interesting?

Network Vendors: Is the assumption correct that a network process, routing process withdrawing a path, most of the time knows why it acts that way, and could potential make this reason and cause information available?

Academia: Would it help if network operators would provide well defined labeled operational and analytical data to enable and validate their research?

Everybody: Should these symptoms be clearly described and standardized for a common terminology so that operators, researchers and anomaly detection systems alike understand their meaning and learn and act accordingly?

Semantic Metadata Annotation for Network Anomaly Detection

Annotate **Operational** Data YANG Module

```
module: ietf-symptom-semantic-metadata
  +--rw symptoms
    +--rw symptom* [id]
      +--rw id          string
      +--rw description      string
      +--rw metrics* [metric]
        | +--rw metric    string
      +--rw start-timeyang:date-and-time
      +--rw end-time   yang:date-and-time
      +--rw concern?  uint8
    +--rw source
      | +--rw (source-type)
      | | +---:(human)
      | | | +--rw human      empty
      | | +---:(algorithm)
      | | | +--rw algorithm  empty
      | | +---rw name?      string
      | +--rw (plane)?
      | | +---:(forwarding-plane)
      | | | +--rw forwarding-plane  empty
      | | +---:(control-plane)
      | | | +--rw control-plane      empty
      | | +---:(management-plane)
      | | | +--rw management-plane  empty
      | +--rw (outlier-type)?
      | | +---:(global)
      | | | +--rw global      empty
      | | +---:(contextual)
      | | | +--rw contextual  empty
      | | +---:(collective)
      | | | +--rw collective  empty
      | | +---:(other)
      | | | +--rw other      empty
      +--rw action?  string
      +--rw reason?  string
      +--rw relation? string
```

- Text

Semantic Metadata Annotation for Network Anomaly Detection

Annotate **Analytical** Data YANG Module

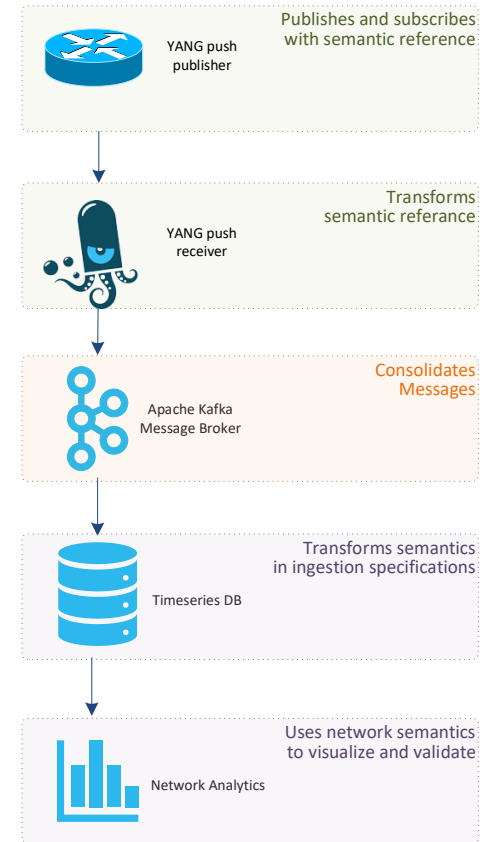
```
module: ietf-incident-label
+--rw incidents
  +--rw incident* [id]
    +--rw id          string
    +--rw description  string
    +--rw start-time   yang:date-and-time
    +--rw end-time     yang:date-and-time
    +--rw concern-score uint8
  +--rw symptoms
    | +--rw symptom* [id]
  +--rw source
    +--rw (source-type)
    | +--:(human)
    | | +--rw human      empty
    | +--:(algorithm)
    | | +--rw algorithm  empty
    +--rw name?         string
```

- **Text**

Semantic Metadata Annotation for Network Anomaly Detection

Next steps

- **Do you realize the gaps and how it could be resolved?**
- This document looks for a community and working group who cares about helping network operators monitoring their networks more efficiently, have interest in Network Anomaly Detection, bridging network and data engineering, operator, vendors and academia, by writing the semantics and ontology of network symptoms for operational and analytical data.
- This work will unveil what is missing in Network Telemetry data and provide input for other documents to enable a more detailed and holistic view from the network.



thomas.graf@swisscom.com
wanting.du@swisscom.com
alex.huang-feng@insa-lyon.fr
vincenzo.riccobene@huawei-partners.com
antonio.roberto@huawei.com

29. October 2023