# CNCF Compliance-grc Use Cases

Hubbert Smith – [hubbert@i4ops.com](mailto:hubbert@i4ops.com)

Slack HubbertSmith https://cloud-native.slack.com/archives/D06PVNSQ5AB

[1 CNCF Compliance-grc Use Cases 1](#_Toc169774054)

[1.1 Mission and Success Criteria 2](#_Toc169774055)

[1.2 DEFINITIONS 4](#_Toc169774056)

[2 Problem Framing 6](#_Toc169774057)

[2.1 Common Data Breaches (spoiler: usually credential abuse) 6](#_Toc169774058)

[3 PERSONAS 7](#_Toc169774059)

[3.1 CISO-office Persona (compliance risk assessment, compliance policy author, policy project oversight) 7](#_Toc169774060)

[3.2 BISO Persona Business Information Security Owner 9](#_Toc169774061)

[3.3 Programmer and Sysops Personas (write compliance as code, use APIs, support/improve) 9](#_Toc169774062)

[3.4 Project Manager (PM) Persona (create workspace, add/rm data, add/rm users. config alarms/alerts/reports) 10](#_Toc169774063)

[3.5 AI/Data Science User Persona 10](#_Toc169774064)

[3.6 3rd Party Credentialed User Persona 10](#_Toc169774065)

[4 Auditing 11](#_Toc169774066)

[5 Compliance, access and data breach 11](#_Toc169774067)

[5.1 Compliance Catalog Risk and Policy 12](#_Toc169774068)

[5.2 Policy-risk-profile and Digital Asset Catalog 12](#_Toc169774069)

[5.3 Risk classification in practice 12](#_Toc169774070)

[5.4 Digital Asset Catalog (locations): where is my (sensitive) data 13](#_Toc169774071)

[5.5 Compliance boundaries (locations): Storage Volumes 13](#_Toc169774072)

[5.6 Digital Asset Catalog API 13](#_Toc169774073)

[6 REFERENCE 14](#_Toc169774074)

[6.1 Data Lifecycle (copies and derivatives of a Digital Asset and GID) 14](#_Toc169774075)

[6.2 [Example] Digital Asset CATALOG based on policy-risk-profile 15](#_Toc169774076)

[6.3 Api 15](#_Toc169774077)

[6.4 NIST sp 800-171 19](#_Toc169774078)

[6.5 Verizon Data Breach 2023 19](#_Toc169774079)

[6.6 Trend Micro definition of data breach 19](#_Toc169774080)

[6.7 Redhat State of Kubernetes Security report 2023 20](#_Toc169774081)

[6.8 Data Clean Rooms 20](#_Toc169774082)

[6.9 CNCF Example Security Use Case 20](#_Toc169774083)

[6.10 CNCF Compliance TAG CHARTER 20](#_Toc169774084)

[6.11 REFERENCE CNCF Compliance TAG RESPONSIBILITIES 21](#_Toc169774085)

[6.12 Appendix – Overflow Topics 22](#_Toc169774086)

Change Log

|  |  |  |
| --- | --- | --- |
| V1 - 2024-May-06 | First draft | Hubbert Smith |
| V2 - 2024-may-13 | Mission, success criteria, definitions, digital asset, storage volume | Hubbert Smith |
| V3 – 2024-may-29 | Clarified Compliance mission, success criteria and Auditing | Hubbert Smith |
| V4-2024-june-20 | Improved compliance and policy definitions. Added reference to SUSTO. Added references to OSCAL. Clarified role of Programmer/Sysops and auditing. Added CIS references | Hubbert Smith |

## Mission and Success Criteria

### Mission –

**Compliance and Policy Objectives**

* As-simple-as-possible deployment to encourage secured innovation, including 3rd parties. Applied compliance and policy starts with cataloging high value and highly sensitive digital assets; not necessarily every digital asset. Compliance and Policy requires a catalog of digital assets with authority/ownership, location, access logging, policy classification of sensitive digital assets. Not all digital assets are sensitive digital assets.
* Compliance-as-Code includes programmatic methods to automate implementation, validation, remediation, monitoring and reporting of compliance requirements across a given organizations technology ecosystem, including 3rd parties.
* Policy-as-code is machine readable, policy rules, classifications, standards, access/copy rules, signoffs and regulations as declarative code to be assigned to specific digital assets and for sensitive digital asset access including 3rd parties and replicas. Policies additionally define required response in the case of unauthorized access, or breach, of sensitive digital assets.  
  Policy-as-code may include penetration-tests, SBOMs, system hardening checks, exfiltration-tests.
* Compliance-as-code for start-of-day security audit to confirm approved handling, access and security. Compliance-as-code includes ongoing security audit to confirm handling, access and security has been maintained during use of sensitive digital asset(s).  
  (example: US Army corps of engineers may share confidential unclassified construction data with a construction contractor; and mid-project, the project the US Army corps of engineers may instruct the contractor to run an audit to confirm the confidential digital assets have been secured/handled in accordance with policy).

### Success criteria –

* **Usable and easily adoptable**
* **Audit, monitoring, and accountability** systems and procedures.
* **Industry Standard Release Candidate.** Endorsed by X adopters and Y industry bodies.
* **Industry standard** for cataloging, classifying, assigning policies for access, replicas, authority/ownership, breach response. Also measuring, monitoring, auditing and logging access (both authorized and unauthorized), data export (both authorized and unauthorized).  
  Accommodating access to sensitive digital assets by distributed organizations and credentialed 3rd parties such as partners, suppliers, consultants and service providers.
* **Reduced Risk:** of sensitive digital assets in decentralized locations: in cloud, replicated regionally distributed for business continuity, or shared with partners/contractors/consultants across various private infrastructure and cloud infrastructures.
* **Reduced Risk**: of more Data is exposed to more humans in business digitization and digital transformation. Sensitive Data is used for business analytics, data science and AI, carries significant compliance risk.

### SCOPE: Data Security Compliance Catalog and Audit

* **Use cases, user personas**, classification and assignment of policy to each sensitive digital asset. This structured approach determines risk probability, risk impact, sensitive data access approval, access monitoring, breach event detection and breach event response. In-house audit of data handling compliance, and audit of partner data handling compliance.
* **Catalog** defining data breach risk domain (data and policy classification), and defines data beach measurement domain (such as a storage system or volume), access policy based on digital asset classification, and response to unauthorized access event.
* **Sensitive Digital Asset Catalog** will assign and store policies to each digital asset based on asset classification
* **Sensitive Digital Asset Catalog** will assign and store authority/ownership, location (where stored/where replicated)
* **Sensitive Digital Asset Catalog** **accommodates singular digital assets with multiple copies/replicas/snapshots in multiple locations** (eg. different cloud instances, different data centers)
* **Sensitive Digital Asset Audit** new sensitive asset location, and access; at start of project to assess risks and controls  
  **Sensitive Digital Asset Audit** mid-project to assess risks, controls, access (both authorized and unauthorized) and export/exfil/copy/replica. Audit will confirm access policy adherence or non-adherence. Sensitive Digital Asset Audit Output confirms sensitive Digital Assets stay within the designated secure volume(location).  
  **Logging for Auditability** requires Authorized users, employees or 3rd parties, may read/write/edit/merge sensitive digital assets within designated data volume (location).  
  **Logging for Auditability** requires logging of user, 3rd party access both authorized and unauthorized. Auditability requires logging of data movement, both authorized and unauthorized.
* **Audit, monitoring, and accountability** systems and procedures.
* **Usability, Adoptability and Deployment**  
  1) Low friction, easy to implement and use monitor for data breaches – including assets accessed by 3rd parties.   
  2) Low friction easy to use methods to identify protected digital assets, classify digital assets, digital asset owners – including assets accessed by 3rd parties.   
  3) For large or small projects, small or large digital assets, private data centers or public cloud, accessed by employees and 3rd party participants.

### OUTSIDE of SCOPE:

* Not an exhaustive Legal compliance/assessment/interpretation – This document is NOT an exhaustive treatment of any legal requirement. It is NOT a representation or claim of legal compliance. It is the responsibility of others to conduct a legal assessment of the varied and changing legal compliance practices for their respective organizations. Legally required regulations include, but are not limited to GDPR, HIPPA, California Consumer Privacy Act (CCPA), California Privacy Rights Act (CPRA).
* Not an exhaustive Government agency compliance assessment/interpretation – This document is inspired by various agencies such as NIST, this is NOT an exhaustive treatment any agency guidelines.
* Not Security Architecture, Computer Architecture, or Data Architecture  
  (these are downstream). Cyber Security: Network Security, Identity access management, Data security. Not Access Control, User Authentication, Least Priv, Identity Access Management, Biometrics, 2 factor auth or Roles based access control. These are beyond the scope of this spec. These are used \_with\_ data sec systems.
* Not Automated Asset discovery, automated classification, or automated digital asset tracking – This assumes BISO (or similar title/responsibility) will start with prior discovery or manual catalog of sensitive digital assets of the respective business unit or operational unit.

## DEFINITIONS

* **OSCAL definitions -** <https://pages.nist.gov/OSCAL-Reference/models/v1.1.2/complete/json-reference/>
* **Assurance Case –** well organized evidence demonstrating some claim about a system is true.
* **Data Breach –** an event when data is taken from authorized data volume to an unauthorized data volume or destination. Data breach includes compliance-relevant data sets; rather than single records.
* **Data Volume –** logical storage unit, providing contiguous access to many physical blocks on many disk drives or many solid state devices.  
  Data Volume examples include an ext4 file system, a ZFS file system, virtual volumes, elastic cloud volumes, storage pools, RAID volumes, data replicas.  
  Data Volume software manages the mapping of physical blocks to filesystem, filesystem hierarchies, filesystem access privileges, journaling, snapshots.
* **Digital Asset** **Classification –** Organization CISO establish appropriate classification categories (not unlike .mil classified Top-Secret, classified Secret, unclassified).   
  Digital asset classification categories define allowed access, approvals for access, approved locations, and required response in the event of unauthorized access or breach of a given digital asset. Each digital asset will carry an assigned category and owner.
* **Digital Asset** **Instance –** one Digital Asset for tracking: a singular, homogeneous, non-public, dynamic data set. Each Digital asset has a unique digital asset identifier and digital asset owner (persona such as CISO or BISO). Data is dynamic, digital assets change with time.   
  A single digital asset, with a unique ID, will have plural copies/derivatives/snapshots of a digital asset instance. Examples include plural secondary copied data sets for analysis, derivative data-sub-sets, copies, replicas, and archives. Digital assets are typically stored in data volumes. Digital assets but may also be stored in block-storage systems, blob/object storage systems or compressed tape or virtual tape archival systems.
* **Digital Asset Location** – physical location of a Digital Asset stored (a datacenter/geography). Digital Asset Location also includes computer or storage systems where data is stored. A digital asset may have plural locations/copies: data backups and snapshots protect current/active data for BC/DR; usually stored in different datacenter/locations.   
  Older data is archived to warm/cold/tape archival storage systems. These plural instances of the same data match our definition of a single Digital Asset instance with plural versions/locations.
* **Digital Asset Owner** – CISO Chief Information Security Officer or BISO Business Information Security Official. Responsible for the use of sensitive data and security of sensitive data.  
  Digital Asset Owners may oversee data operations of an entire organizations or business entities,   
  or oversee data operations of lines of business with P&L responsibility,  
  or oversee data operations of functional areas such as Payroll, AR/AP, Help Desk, Sales, Supply Chain Operations, or Quality Assurance.
* **Digital Asset sharing** - Policy typically creates a new Digital Asset identifier when data is sent to partner. Policy typically adds contract terms for data handling, data access, data destruction, data security/RBAC, alerting in case of breach, and audit.
* **Enterprise Environment** – pre-existing systems and cybersecurity safeguards such as Two factor authentication, Identity access management, Roles based access controls, Workspaces (see def.).
* **Exfiltration** – unauthorized data handling, unauthorized data transmission, unauthorized data movement or data storage.
* **Exfiltration Audit** – programmatically attempt unauthorized exfiltration of dummy test data from a secured digital asset location (in production or pre-production), run multiple programs to attempt exfiltration, audit tests for cp, SFTP, SCP, etc etc.  
  Results logged shall show “exfiltration prevented” or “exfiltration allowed” for each audit test
* **Export Audit**– programmatically confirm authorized export of dummy test data from a secured digital asset location (in production or pre-production) to a defined destination.   
  Export data is defined as specific digital asset, specific user/process send to specific destination.   
  Results logged shall show “authorized export allowed” or “authorized export prevented”.
* **Export** – authorized data handling, authorized data transmission or authorized data movement or data storage. Typically, pre-authorized by CISO, BISO, or PM and typically logged.
* **Digital Workspaces** – any place data may be accessed, edited, read, or transferred to-from.   
  Workspaces may include individual PCs, Servers/Clusters in private data centers, Co-location data centers, Cloud systems. Workspaces are Increasingly cloud-based and/or private-cloud. Workspaces use a variety of infrastructure system, virtual machine or container technologies. This document uses kubernetes terminology, but this standard does NOT mandate kubernetes.
* **FTE** – full time employee of the data-owning organization.
* **Policy-as-code and Compliance-as-code** each sensitive digital asset is assigned a policy.  
  Policies define desired security state, typically includes approved locations/access/signoffs, penetration-testing, SBOMs, system hardening checking, exfiltration-testing requirements. **Compliance-as-code** confirms each sensitive digital asset has been handled in accordance with assigned policy. Compliance as code compares actual locations/accessed-by/signoffs/activity logs/access logs to the desired security state as defined by policy assigned to the digital asset(s)
* **Third parties** – any non-FTE individual, contractor, consultant, team or business accessing data

# Problem Framing

**Data Breach is STILL daily news …** (Why?) Humans touching data and credential abuse is a frequent source of breach. Departing employees, contractors, credential abusers sysadmin abuse, sloppy credentials. NDAs, training, MSSP monitoring are not effective to achieve compliance.  
**More strangers touching more data leads to data breach.**

**Workplaces** - increasingly dispersed offices, remote employees, remote contractors and remote partners/service providers.

**People**: Employee churn is common. Global Tech Talent, Project Consultants, Contract workers with task-driven data access, sys-admin, cloud-admin, and storage-admin privs are common.

**Business Transformation** - AI/Data Sci digital transformation projects involve large, sensitive data sets which are subject to GDPR, etc. Additionally, AI/Data Sci projects involve business-sensitive and unclassified government-secret.

**Data breach** – is the act of taking data from authorized data volume to an unauthorized data volume. Some data is highly secured with tightly defined authorized data volumes. Other data is does not meet the risk classification thresholds, may be stored on any data volume.

**Accountability** – CISO and BISO are accountable for securing sensitive digital assets. CISO and BISO are also accountable for Auditing and Monitoring secure data volumes where sensitive digital assets are stored.   
Digital Assets include primary/transactional data, replicas, secondary analytics data, derivative data from analytics and business continuity backups/snapshots.

## Common Data Breaches (spoiler: usually credential abuse)

* Insider threats - Whenever humans touch data, there is risk. Credential abuse is typically the root of unauthorized exfiltration. Data breaches, cited in daily news, involve credentialed employees, credentialed contractors, lax account/password care, and syspriv abuse.
* Stolen Credentials – use login credentials, often exposed by other breaches
* Phishing attacks – send fraudulent emails to employees to trick them into revealing login credentials or installing malware (or both)
* Malware, Ransomware – encrypt data and hold hostage until ransom is paid
* Code injection - Hackers inject malicious code into government websites or databases to steal data. SQL injection and cross-site scripting (XSS) are common techniques
* Supply chain attacks - Targeting third-party vendors to access sensitive data; usually credential abuse. When data is shared with any partner, that data is at risk. Risk triage will categorize the probability and the consequences, and contractually establish required data handling, and audit, with partner.

# PERSONAS

1. **CISO-office Persona:** Compliance risk assessment, compliance policy author, policy project oversight
2. **BISO Persona:** Business Information Security Owner persona
3. **Programmer and Sysops Personas:** Write compliance as code, use APIs, support/improve
4. **Project Manager Persona:** create workspace, add/rm data, add/rm users. config alarms/alerts/reports
5. **AI/DataSci User Personas** Data user, data scientist, knowledge worker
6. **3rd parties**

|  |  |
| --- | --- |
| CISO-office Persona (compliance risk assessment, compliance policy author, policy project oversight) | |
| Compliance Policy Risk assessment | As a CISO, I am responsible for establishing, managing, communicating, and monitoring Compliance Policy – risk categories, data location/replication, data access criteria, response criteria, discovery, event root cause/corrective action.  **As a CISO, I develop, use and oversee:**  **Digital data asset catalog** – collect sensitive digital assets into a managed list/database and tag each digital asset w/ a risk category.  Data catalog must include derivative data, snapshots, replicas, archives  **Risk Categories**: risk tier 1,2,3,4 risk scenario(s), probability, impact Define a digital asset deletion policy for each category  **Location/replication:** Systems and software/data handling catalog Cataloged systems and software is used in compliance data access (database SW, explorers, BI tools, AI/DataSci tools, data replication/snap/backup/archival)  **Data Access Policy:** Definition of who can access data: employee persona, contractor persona, partner persona.  **Response Policy**: Clearly defined response criteria in the case of tier1 breach. Clearly defined response in the case of a Tier2 breach … etc.  **Discovery**: ongoing discovery of data which carries legal risk:  GDPR: Name, DOB, physical, physiological, genetic, mental, economic info ID-num/SSN#, location, online identifiers (eg. Equifax, Microsoft, Marriott, Cathay Pacific, British Air) HIPAA: personal identifiers, medical records/tests/exams/treatments, billing/payments, appointments (eg. MCNA breach, PharMerica, Community health systems breach 2023)  **Discovery**: ongoing discovery of data which carries business risk:  SEC requires filings of Breach incidents which potentially impact stock values of Publicly traded companies. such as Demand/Supply/Ops data: Sales, Supply Chain, Help Desk, IP, IT infrastructure (Okta, 23andMe, Tesla  **Discovery**: of data which carries Public Security Risk: U.S. Gov, NATO, EU classified documents, data collection and monitoring, military intelligence, covert intelligence (eg Snowden, Hanssen)  AI-specific risk assessment balancing the need for AI training with the risk of data breach; including LLM queries which can include sensitive data. |
| Compliance policy author, knowledge transfer, change manager | As a CISO, I am responsible for delivering Knowledge-transfer to employees, contractors and partners Catching people doing the right thing is essential Private guidance when people do the wrong thing, is essential |
| Event Response | As a CISO, I am responsible for developing, dry-running a Response plan for each risk category. And when breach is detected; I am responsible for checking the risk category tag, and executing the response plan. |
| Policy for Partner data sharing | As a CISO, I establish policy of data sharing.  I establish a Standardized Partner contract for partners who receive sensitive data. The contract covers required data security, roles based access, breach monitoring/alerting. |
| Policy monitoring and oversight | As CISO, I am responsible for regular review of Compliance reports and telematics produced by workspaces and (if needed) take remedial action.  As CISO, I am responsible for accurately logging every breach, with risk category tag, with root cause & corrective action |

|  |  |
| --- | --- |
| BISO Persona Business Information Security Owner | |
| Compliance Policy Risk assessment | As BISO, I am responsible for domain-specific data Risk Assessment. I assess impact of possible data breach data-set, by data-set within my domain. |
| 3rd party Risk Management | As BISO, I am responsible for domain-specific data handled by contractors, consultants, partners, vendors.  Including Business Services such as Book keeping, AR/AP, Shipping, Help Desk, Call center, AI projects. And vendors providing data services such as data storage, backup, bc/dr, Sys-admin, Cloud-Admin, Storage-admin services. Vendors may also provide AI/DataSci services, tagging, data processing, and consultants, LLMs |
| Compliance policy author, knowledge transfer, change manager | As a BISO, I am responsible for domain-specific Data Identification, Data Classification, Data Access policy. I am also responsible for setting/executing Compliance monitoring criteria. And, setting/executing change management (new policy, new approvals, new digital assets, new partner-data-access, partner-contract)  Domain-Specific Knowledge-transfer is essential for adoption Catching people doing the right thing is essential Private guidance when people do the wrong thing, is essential |
| Policy project definition and oversight | As a BISO, I approve data access for domain-specific business transformation projects. I ensure that the Project success criteria includes conformance, knowledge exchange, with programmers, AI/DataSci users, 3rd parties, and end users |
| Policy monitoring and oversight | As BISO, I am responsible for date used by domain-specific Projects and Workspaces. I approve adding/revoking user access, adding/deleting data,  reporting and telematics data collection and records keeping.  As BISO, I am alerted when breach is detected, check the risk category tag, response plan and respond accordingly. I am responsible for after-event root cause and corrective action |

|  |  |
| --- | --- |
| Programmer and Sysops Personas  (write compliance as code, use APIs, support/improve) | |
| Compliance as code | As Programmer, I develop code based on specifications and feature lists provided by business personas. I write, debug, improve, and audit compliance-related code Perform QA, SBOMs, Audits, Pen-Tests and Exfil-Tests. I update Digital Asset catalog. |
| Catalog sensitive digital assets | As Programmer/Sysops, I will manage digital asset locations, copies/replicas, roles based access and logging. These will be recorded in the catalog (actions table)  As Programmer/Sysops, I will establish repeatable steps to add/edit compliance and policy catalog.  I will establish repeatable steps to audit digital workspaces at the time of data-driven project creation; to confirm appropriate security is in place prior to adding/accessing sensitive digital asset to the project. I will establish repeatable steps to audit digital workspaces, access logging, data logging as compared to policy catalog; to confirm policy adherence of sensitive digital access handing. |
| Compliance API, knowledge transfer, change manager | As programmer, I am responsible for domain-specific/project-specific knowledge exchange to Project managers, SYSops, Data Science users and BISOs via classroom, video, user guide, as appropriate. |

|  |  |
| --- | --- |
| Project Manager (PM) Persona (create workspace, add/rm data, add/rm users. config alarms/alerts/reports) | |
| Create workspace (think K8s) | As PM, establish workspace, invite users, upload digital assets, connect to digital assets in place. Allocate compute, storage, network infrastructure. |
| Workspace  Users, data, alerts | As PM I manage data access whitelist (add user, remove user). I manage export whitelist, and Workspace/Project-specific alarms/alerts |

|  |  |
| --- | --- |
| AI/Data Science User Persona | |
| AI/Data Sci credentialed users | As AI/DataSci user, I build AI/Data Sci projects. As AI/DataSci user, I interact with partners, contractors, consultants performing AI/Data Sci project  As a Data Scientist, I perform initial data quality assessment, data cleansing, cross referencing multiple data sets. I implement publish-subscribe in-bound data and out-bound data processes I perform AI model training, AI model tuning, AI model improvement I deliver AI inference systems, independent of AI training systems I perform ongoing AI inference processing, fresh data updates, ongoing data cleansing/merging, drift and bias correction. I create Business Dashboards, reports, I deliver ongoing dashboards, reports |

|  |  |
| --- | --- |
| 3rd Party Credentialed User Persona | |
| 3rd party credentialed user | As a 3rd party credentialed user participating in projects with sensitive digital assets from owner; I access sensitive digital assets owned by others during performance of work similar to Data Scientist personas, above.  Additionally, as a 3rd party-credentialed user, I respond to requests for compliance/policy audits as required for the owner of the sensitive digital assets under my control. |

# Auditing

**Audit for authorized Export or Access:** programmatically confirm authorized export of dummy test data from a secured digital asset location (in production or pre-production) to a defined destination. Export data is defined as specific digital asset, specific user/process send to specific destination.   
Results logged shall show “authorized export allowed” or “authorized export prevented”.

**Start of day Audit for Unauthorized Exfil or Access:** a programmatic test which attempts unauthorized exfiltration of dummy test data from a secured digital asset location (in production or pre-production), run multiple programs to attempt exfiltration, audit tests for cp, SFTP, SCP, etc etc.  
Results logged shall show “exfiltration prevented” or “exfiltration allowed” for each audit test.

**Audit for ongoing unauthorized exfil or access:** identify digital asset(s), identify classification and owner, collect activity logs, compare with catalog and policy.

**Audit (dry run) Compliance Policy response in the Event of a breach:** immediately show what digital assets are breached, classifications, breach response based on classification and owners – including assets accessed by 3rd parties. Communicate to designated Digital Asset owners and executive team (legal, investor relations, finance, operations, affected 3rd parties, partners, customers, and/or suppliers for the specific digital asset breached.

# Compliance, access and data breach

* **Compliance and first principles:** Prevent compliance violations: prevent unauthorized access or data breach of compliance policy-relevant sensitive data.
* **Why Measure:** The industry lacks a standard means to measure and audit policy-relevant sensitive data. Self-audits of sensitive data handling and partner/suppler/consultant audits of sensitive data handling are relevant.
* **What to measure**: sensitive data listed in Digital Asset Catalog.
* **Where to measure(example)**: storage volume(s) storing policy-relevant sensitive data.   
  The storage volume is the perimeter, keep the data IN.
* **How to measure**: policy-relevant access and copy of digital assets. Requires policy, asset GID, storage volumes are equipped with API, the API logs and updates digital asset catalog of updates.
* **System makeup**: constantly changing data, cataloged digital assets, asset policies (in catalog), asset access (in catalog). Locations (storage volumes, in catalog), digital asset GID (in metadata), export/exfil detection in (data volume). Logging and catalog updates of both authorized exports and unauthorized exfiltration.
* **Focus on high risk policy-relevant sensitive data.** Waste no time on low-risk, non-sensitive data.

## Compliance Catalog Risk and Policy

This defines **policy-relevant sensitive digital assets**, assigns policy with digital asset classification, tracks access, copies, replicas. The Digital Asset Catalog is the basis of auditability.

* Data owners identify policy-relevant digital assets and update the Digital Asset catalog.
* Data storage for policy-relevant digital assets will have API for logging of authorized export and unauthorized exfil.
* Data Storage API will capture approved data movement (export), or unapproved data movement (exfiltration). Data Storage API captures both local lateral data movement or network data movement.
  + Logging will track local lateral data movement: from approved data storage to some other local data storage (either approved or unapproved).
  + Logging will track network data movement: from approved data storage to some other networked system (either approved or unapproved).
* API (for example, but not limited to OSCAL) for policy-relevant data on data storage systems will send event data to digital asset catalog.

## Policy-risk-profile and Digital Asset Catalog

This section seeks to apply existing Digital Asset Catalogs to Policy/Risk to Data breach prevention, monitoring, audit. Catalog is populated by each digital asset subject to policy.  
a suggested data structure is included in the appendix

**Risk Class 1** – Type of data: digital asset breach directly results in significant harm to the business. Compliance violations, legal consequences, trade consequences, brand reputation and stock value consequences. Public sector – Government classified secret, unclassified secret.  
**Policy for Risk Class 1**: required actions in the case of breach or audit. Who is alerted (CEO, CIO, CISO, Legal, CFO), forensics information, root-cause-corrective-action required

**Risk Class 2** – digital asset breach results in some harm to the business.   
public sector – Government unclassified not secret  
**Policy for Risk Class 2**: required actions in the case of breach or audit. Who is alerted (CISO, BISO Legal, CFO), forensics information, root-cause-corrective-action required

**Risk Class 3** – digital assets posing little or no harm to business.   
**Policy for Risk Class 3**: required actions in the case of breach or audit. Who is alerted (CISO, BISO), forensics information, root-cause-corrective-action required

## Risk classification in practice

* In Data catalog, each digital asset is assigned a risk class, digital asset global-id, data owner
* When data is breached, Risk classifications deliver clear policy and actions required; by specific roles assigned to specific individuals.
* When data sharing is requested, Risk classifications serve as clear policy for sharing specific data to specific FTE roles and credentialed 3rd parties.
* When audits are required, Risk classifications confirm policy compliance, for sharing specific data to specific FTE roles and credentialed 3rd parties.
* Organizations are free to establish Risk Classes to fit organizational needs and policy needs.

## Digital Asset Catalog (locations): where is my (sensitive) data

This specifies auditable minimum baseline capabilities for data security and privacy auditability.   
This does not imply or specify specific databases, data catalogs or similar, just the auditable baseline capabilities. This Digital Asset Catalog with policies and tracking is essential for compliance.

**Table ASSETS** - AssetID, Asset name, Asset business owner, Asset policy/risk class

**Table POLICIES** - Policy, risk-classification, alert, action, access allowed policy, access prohibited policy

**Table Asset ACCESS** - Name, AssetID, Copy event, date, location, accessed by(people)

* Catalog of Digital **assets subject to security or privacy policy** (just the sensitive stuff). Catalog assigns policy-risk classification: allowed access, disallowed access, breach-event response
* Unique ID per digital asset.
* API+catalog tracks copies/derivatives/snapshots of Sensitive Digital Assets (see reference-data lifecycle).
* API+catalog tracks access granted to Sensitive digital assets and actual access of digital access by humans, FTE, and credentialed 3rd parties.
* Actions from APIs are aggregated to the Digital Asset Catalog.

## Compliance boundaries (locations): Storage Volumes

* Sensitive Digital Asset Security dictates some sort of boundary (aka secure location).   
  Such as defined authorized data volume with classification-appropriate exfil security.
* Data breach (aka exfiltration) is when data is taken from authorized secure location to an unauthorized (elsewhere).

## Digital Asset Catalog API

This specifies auditable minimum baseline capabilities for data security and privacy auditability.   
This does not imply or specify specific databases, data catalogs or similar, just the auditable baseline capabilities. This Digital Asset Catalog with policies and tracking is essential for compliance.

**Table ASSETS** - AssetID, Asset name, Asset business owner, Asset policy/risk class

**Table POLICIES** - Policy, risk-classification, alert, action, access allowed policy, access prohibited policy

**Table Asset ACCESS** - Name, AssetID, Copy event, date, location, accessed by(people)

This API populates the Digital Asset Catalog;  
It is significant for security and privacy compliance;   
whenever the digital asset is added/copied;  
when access is granted/revoked;  
when digital assets are added/changed/exported;  
when digital assets experience unauthorized egress attempt

Add digital asset to catalog – add name, add owner, add policy class, auto-add global-ID, point to file location.

Copy existing digital asset – preserve global-id, point to new file location, date stamp, UID/PID doing the copy.

Add access permission to digital asset.

Add export/exfil api to secure data storage (data volume).

# REFERENCE

**BBVA/susto OSS Open Security Control Testing at Scale.pdf**

<https://github.com/BBVA/susto/blob/master/docs/docs/OSS-Open%20Security%20Control%20Testing%20at%20Scale.pdf>

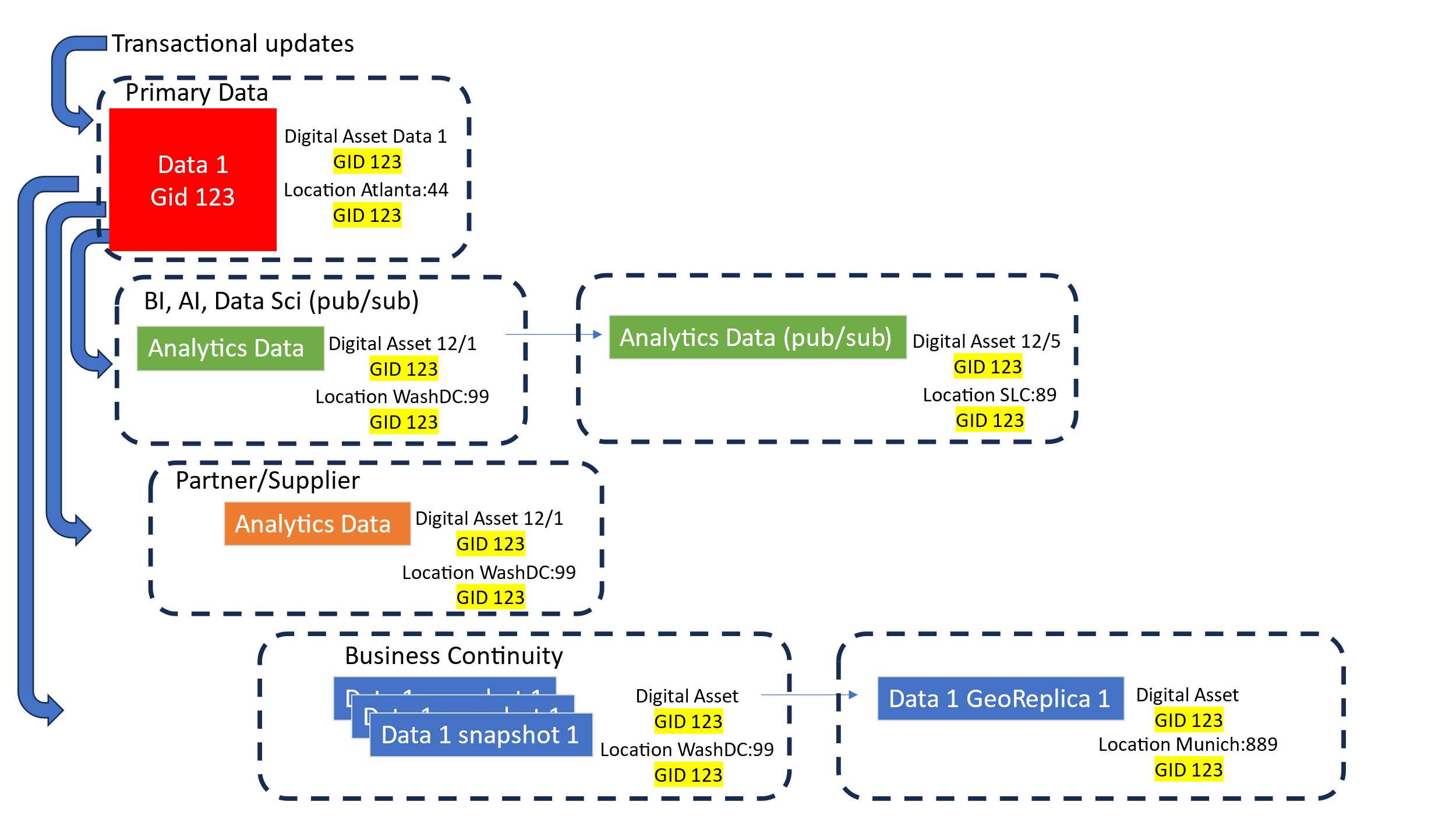
**Catalog of Sensitive Data**

Digital **Assets** (in a catalog)

Digital Asset **Policies** for each digital asset: locations, permitted access, action if breached, owner

Digital Asset **Access** monitor, log, access, export and exfiltration: what Digital Asset, who, when, location

## Data Lifecycle (copies and derivatives of a Digital Asset and GID)



## [Example] Digital Asset CATALOG based on policy-risk-profile

This Digital Asset Catalog with policies and tracking is essential for compliance.  
This outlines minimum required functions and tracking.

**Table ASSETS** - AssetID, Asset name, Asset business owner, Asset policy/risk class

**Table POLICIES** - Policy, risk-classification, alert, action, access allowed policy, access prohibited policy

**Table AsssetACCESS** - Name, AssetID, Copy event, date, location, accessed by(people)

-- Table 1: Assets

CREATE TABLE Assets (

AssetID INT PRIMARY KEY,

AssetName VARCHAR(255) NOT NULL,

AssetBusinessOwner VARCHAR(255) NOT NULL,

AssetPolicyRiskClass VARCHAR(50) NOT NULL

);

-- Table 2: Policies

CREATE TABLE Policies (

Policy VARCHAR(50) PRIMARY KEY,

RiskClassification VARCHAR(50) NOT NULL,

Alert VARCHAR(255) NOT NULL,

Action VARCHAR(255) NOT NULL,

AccessAllowedPolicy VARCHAR(255) NOT NULL,

AccessProhibitedPolicy VARCHAR(255) NOT NULL

);

-- Table 3: AssetAccess

CREATE TABLE AssetAccess (

Name VARCHAR(255) NOT NULL,

AssetID INT NOT NULL,

CopyEvent VARCHAR(50) NOT NULL,

AccessDate DATE NOT NULL,

Location VARCHAR(255) NOT NULL,

AccessedBy VARCHAR(255) NOT NULL,

FOREIGN KEY (AssetID) REFERENCES Assets(AssetID)

);

## Api

The API is essential populates/updates the Digital Asset Catalog   
Eg. add new Digital assets, assign policy, access/location tracking of copies and derivatives  
monitor, alert, audit

### API Add to asset table

from flask import Flask, request, jsonify

import sqlite3

app = Flask(\_\_name\_\_)

# Connect to the SQLite database

conn = sqlite3.connect('assets.db')

c = conn.cursor()

# Create the Assets table if it doesn't exist

c.execute('''CREATE TABLE IF NOT EXISTS Assets

(AssetID INTEGER PRIMARY KEY, AssetName TEXT NOT NULL, AssetBusinessOwner TEXT NOT NULL, AssetPolicyRiskClass TEXT NOT NULL)''')

@app.route('/assets', methods=['POST'])

def add\_asset():

# Get the asset data from the request

asset\_data = request.get\_json()

# Extract the asset details

asset\_name = asset\_data['AssetName']

asset\_business\_owner = asset\_data['AssetBusinessOwner']

asset\_policy\_risk\_class = asset\_data['AssetPolicyRiskClass']

# Insert the asset into the database

c.execute("INSERT INTO Assets (AssetName, AssetBusinessOwner, AssetPolicyRiskClass) VALUES (?, ?, ?)",

(asset\_name, asset\_business\_owner, asset\_policy\_risk\_class))

conn.commit()

# Get the ID of the newly inserted asset

asset\_id = c.lastrowid

# Return the asset details with the assigned ID

return jsonify({'AssetID': asset\_id, 'AssetName': asset\_name, 'AssetBusinessOwner': asset\_business\_owner, 'AssetPolicyRiskClass': asset\_policy\_risk\_class}), 201

if \_\_name\_\_ == '\_\_main\_\_':

app.run(debug=True)

send a POST request to http://localhost:5000/assets with a JSON payload containing the asset details, like this:

json

{

"AssetName": "Confidential Documents",

"AssetBusinessOwner": "Finance Department",

"AssetPolicyRiskClass": "Confidential"

}

### API add to Policies table

from flask import Flask, request, jsonify

import sqlite3

app = Flask(\_\_name\_\_)

# Connect to the SQLite database

conn = sqlite3.connect('assets.db')

c = conn.cursor()

# Create the Policies table if it doesn't exist

c.execute('''CREATE TABLE IF NOT EXISTS Policies

(Policy TEXT PRIMARY KEY, RiskClassification TEXT NOT NULL, Alert TEXT NOT NULL, Action TEXT NOT NULL, AccessAllowedPolicy TEXT NOT NULL, AccessProhibitedPolicy TEXT NOT NULL)''')

@app.route('/policies', methods=['POST'])

def add\_policy():

# Get the policy data from the request

policy\_data = request.get\_json()

# Extract the policy details

policy = policy\_data['Policy']

risk\_classification = policy\_data['RiskClassification']

alert = policy\_data['Alert']

action = policy\_data['Action']

access\_allowed\_policy = policy\_data['AccessAllowedPolicy']

access\_prohibited\_policy = policy\_data['AccessProhibitedPolicy']

# Insert the policy into the database

c.execute("INSERT INTO Policies (Policy, RiskClassification, Alert, Action, AccessAllowedPolicy, AccessProhibitedPolicy) VALUES (?, ?, ?, ?, ?, ?)",

(policy, risk\_classification, alert, action, access\_allowed\_policy, access\_prohibited\_policy))

conn.commit()

# Return the policy details

return jsonify({'Policy': policy, 'RiskClassification': risk\_classification, 'Alert': alert, 'Action': action, 'AccessAllowedPolicy': access\_allowed\_policy, 'AccessProhibitedPolicy': access\_prohibited\_policy}), 201

if \_\_name\_\_ == '\_\_main\_\_':

app.run(debug=True)

send a POST request to http://localhost:5000/policies with a JSON payload containing the policy details, like this:

json payload

{

"Policy": "ConfidentialDataPolicy",

"RiskClassification": "High",

"Alert": "Confidential data accessed",

"Action": "Notify data owner and security team",

"AccessAllowedPolicy": "Authorized personnel only",

"AccessProhibitedPolicy": "No external access allowed"

}

### API add to access table

from flask import Flask, request, jsonify

import sqlite3

from datetime import datetime

app = Flask(\_\_name\_\_)

# Connect to the SQLite database

conn = sqlite3.connect('assets.db')

c = conn.cursor()

# Create the AssetAccess table if it doesn't exist

c.execute('''CREATE TABLE IF NOT EXISTS AssetAccess

(Name TEXT NOT NULL, AssetID INTEGER NOT NULL, CopyEvent TEXT NOT NULL, AccessDate TEXT NOT NULL, Location TEXT NOT NULL, AccessedBy TEXT NOT NULL, FOREIGN KEY (AssetID) REFERENCES Assets(AssetID))''')

@app.route('/asset-access', methods=['POST'])

def add\_asset\_access():

# Get the asset access data from the request

access\_data = request.get\_json()

# Extract the asset access details

name = access\_data['Name']

asset\_id = access\_data['AssetID']

copy\_event = access\_data['CopyEvent']

access\_date = datetime.now().strftime('%Y-%m-%d %H:%M:%S')

location = access\_data['Location']

accessed\_by = access\_data['AccessedBy']

# Insert the asset access into the database

c.execute("INSERT INTO AssetAccess (Name, AssetID, CopyEvent, AccessDate, Location, AccessedBy) VALUES (?, ?, ?, ?, ?, ?)",

(name, asset\_id, copy\_event, access\_date, location, accessed\_by))

conn.commit()

# Return the asset access

send a POST request to http://localhost:5000/asset-access with a JSON payload containing the asset access details, like this:

Json payload

{

"Name": "John Doe",

"AssetID": 1,

"CopyEvent": "Copy",

"Location": "Office",

"AccessedBy": "John Doe"

}

### API data export/exfil

Data storage for policy-relevant data will have logging of authorized export and unauthorized exfil.

Data movement is either approved (export), or unapproved (exfiltration).  
Data movement is either local lateral data movement or network data movement.

Logging will track local lateral data movement: from approved data storage to some other local data storage (either approved or unapproved)

Logging will track network data movement: from approved data storage to some other networked system (either approved or unapproved).

API for policy-relevant data on data storage systems will log lateral data movements or network data movement events, and send event data to digital asset catalog – assetAccess.

### API policy audit

A report, run against the Digital Asset Catalog.   
Policy table assigns policy to each digital asset. Policy defines which personas can access which data.   
Access table shows both authorizations and actual access, exports, exfiltration.  
Audit reports will flag policy violations and data escapes, by digital asset global ID

## NIST sp 800-171

<https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-171Ar3.pdf>

Assessing Security Requirements for Controlled Unclassified Information. Federal perspective and non-Federal perspective

* **Alignment with Agency objectives: NIST SP 800-171Ar3** The security assessment process gathers information and produces evidence to determine theeffectiveness of security requirements by:1) Identifying potential problems or shortfalls in security and risk management programs.  
   >> RISK: data exfiltration involving valid credentials2) Identifying security weaknesses and deficiencies in systems and the environments inwhich those systems operate.  
   >> RISK: data exfiltration involving valid credentials  
  3) Prioritizing risk mitigation decisions and activities.  
   >> AUDIT confirming digital assets are secured from data exfiltration involving valid credentials  
  4) Confirming that identified security weaknesses and deficiencies in the system and environment of operation have been addressed.  
   >> AUDIT confirming digital assets are secured from data exfiltration involving valid credentials  
  5) Supporting continuous monitoring activities and providing information security situational awareness.  
   >> MONITORING/LOGGING attempts to exfiltrate digital assets, involving valid credentials, both successful and unsuccessful.

## Verizon Data Breach 2023

<https://www.verizon.com/business/resources/reports/dbir/>

Key Points

* More likely: Human Element, Malicious insiders, stolen credentials
* Less likely: physical breaches at data center

## Trend Micro definition of data breach

<https://www.trendmicro.com/vinfo/us/security/definition/data-breach>

key points:

* Insider leak: A trusted individual or person of authority with access privileges steals data.
* Payment card fraud: Payment card data is stolen using physical skimming devices.
* Loss or theft: Portable drives, laptops, office computers, files, and other physical properties are lost or stolen.
* Unintended disclosure: Through mistakes or negligence, sensitive data is exposed.

## Redhat State of Kubernetes Security report 2023

<https://www.redhat.com/rhdc/managed-files/cl-state-kubernetes-security-report-262667-202304-en.pdf>

Key Points

* Security is highly de-centralized for containers and orchestrators.
* Safeguarding containerized applications requires different roles to own a piece of the system and processes used in the development life cycle

## Data Clean Rooms

* Data clean rooms are intended for read-access to secured data by humans
* Data clean rooms do not allow write/edit access to secured data, merge multiple datasets or export of a work product
* Therefore, data clean rooms are not useful for data science or AI

## CNCF Example Security Use Case

<https://github.com/cncf/tag-security/tree/main/usecase-personas>

## CNCF Compliance TAG CHARTER

CNCF Compliance TAG charter  
<https://docs.google.com/document/d/1z9xvt-Z97j4CtEH1-nR9sMWul7jQkUi_fNY7BdMPgxM/edit#heading=h.7opelk7xdnar>

Cloud Native systems represent a paradigm shift in both technical and human operations workflows.  The community (and industry) has invested significant time researching and solutioning approaches to Cloud native security concerns and topics: software vulnerabilities, risk management, software component dependencies and infrastructure as code (GitOps), supply chain provenance, malicious attackers, threat models, and technical security assessments.  At the same time, many commercial, non-profit foundations, community and government organizations performing services or providing data storage must abide by national, regional, or local laws and regulations regarding user privacy and data, with assurance of protection of their compute and data processing integrity and resilience. These cross cutting concerns span not only specific technical configuration of software and systems, but also require complex orchestration of human administrative, operational, and design activities, especially when involving  audit activities expecting concrete, reviewable independent audit artifacts.

The motivation of the Compliance WG is cross-disciplinary and focused on bridging purely technical issues to broader legal and regulatory workflows: not only to strive for the prevention of system breaches, but also considering supply chain, operators, data and AI failures while simultaneously considering auditability, non-repudiation, legally required forensic evidence, etc. - across all the various activities required of cloud native operators of all sizes. The Compliance WG plans to curate vendor neutral tools for evidence collection, chain-of-custody in audits, as well as automated workflows for continuous compliance authoring and assessment.

The key areas of the Compliance WG include:

* Building a knowledge base (in GitHub and possibly other tools) and case studies on the How, What, Why and When of operating a cloud native environment within the requirements of legal and regulatory entities that govern clouds, specific industries, and more generally data and public/consumer usage. These requirements are often NOT JUST technical security concerns. Compliance activities and requirements span human activities and performance, system availability and reliability, the combined human and technical aspects of continuity of operations, defining and monitoring data location as well as sovereignty and provenance of the regulated environment components and data.
* Generating specific examples of compliance as code, normalized templates, and tools for automating these both technical and non-technical requirements, control assessment, data analysis, audit and compliance remediation workflows that specifically benefit CNCF projects and their community of users.
* Reviewing industry and governmental standards - eg NIST, PCI, HIPAA, etc - from a cloud native perspective and serving as Subject Matter Experts in the CNCF community for how projects should implement and support these compliance-specific requirements as first class citizen to enable broad adoption of the best practices by commercial, non-profit, governmental, and humanitarian organizations.

## REFERENCE CNCF Compliance TAG RESPONSIBILITIES

* Users/personas/needs/customer demands for industry and regulatory compliance (both human and technical)
* Identifications of areas of focus e.g. human workflows, automated workflows, analytical tools, audit and assessment tools, technical security controls that cut across components and systems and clouds, etc
* Framework for evaluation, audit and reporting - how do products and tools demonstrate compliance?
* Training and automation - what is missing, what is difficult to understand, what knowledge gaps are there? -- prevent exfil
* Work on integrating common tooling across different projects, particularly where that tooling is a CNCF project (but the targets may not be)
* Cross project focus on the projects and efforts the CNCF is funding, helping projects identify needs and providing subject matters to assist
* Recommendations of integrating security tooling with compliance tooling and processes - making both the synergies and unique separations of concern explicit and achieving community consensus.
* Growing CNCF external relationships with interested parties, e.g. NIST and other compliance standards bodies such as FINOS, OSCAL, OpenSSF

## Appendix – Overflow Topics

### Alert/avoid compliance non-conformity

Auditing, SBOM, CIS benchmarks

* Control plane
* KubeVirt virtual machines, and virtual images
* Docker and Podman containers
* KVM
* Bare Metal

Auditing, SBOM

### Socio-Technical

* Harm to people
* Harm to organizations
* Harm to broader ecosystems

### AI Frameworks

NIST AIRMF

ISO 42,001

Ueai framework

### Multi-Organizational data

Shared with Business Partners

Shared for federated Learning and data brokering

Shared for Industry consortiums