COMMONWEALTH of VIRGINIA

Office of the Governor

Lyn McDermid
Secretary of Administration

Virginia Data Governance Council

Meeting Minutes

May 21st, 2024

East Reading Room – Patrick Henry Building

1111 East Broad Street

Richmond, Virginia 23219

1:00 p.m. – 3:00 p.m.

Virtual Attendance Link: https://teams.microsoft.com/l/meetup-join/19%3ameeting_OWE4ODI5YmUtOWE4MC00N2ZhLTg2ZjUtZmQwYzI3Y2FiODVi%40thread.v2/0?context=%7b%22Tid%22%3a%22620ae5a9-4ec1-4fa0-8641-5d9f386c7309%22%2c%22Oid%22%3a%228b5e1e08-5841-4e1e-b617-8592886537f7%22%7d

Member Attendees:

Marcus Thornton, *Office of Data Governance and Analytics*

Rich Rosendahl, *Department of Medical Assistance Services*

Mitzi Fletcher, *Department of Social Services*

TJ Claiborne, *Department of Conservation and Recreation*

Paulose Poovathurkaran, *Department of Behavioral Health and Developmental Services*

Anup Srikumar, *Virginia Department of Health – virtual attendee*

Keon Turner, *Virginia State Police – virtual attendee*

Karen Smith, *Virginia Employment Commission – virtual attendee*

Agenda:

Welcome and Opening Remarks
Roll Call
Data Governance Council Order of Business

NIEM Overview
Cross-State Report on Data Governance and Literacy
Break
Big ID Demo
Informatica Data Quality Demo
ODGA Update
Member and Public Comment
Closing Remarks
Adjourn

Motions:

Motion 1: Approval of Previous Data Governance Council Meeting Minutes (October 2023).

Presenter: Deputy Chief Data Officer, Marcus Thornton

Discussion: Approval of meeting minutes were discussed.

Motion Accepted: Motion moved by TJ Claiborne and seconded by Rich Rosendahl. Motion carried by unanimous vote.

Motion 2: Approval of Data Governance Policy

Presenter: Deputy Chief Data Officer, Marcus Thornton

Discussion: Approval of the Data Governance Policy was discussed.

Motion Accepted: Motion moved by Rich Rosendahl and seconded by TJ Claiborne.

Agenda

Agenda Item 1: Welcome and Opening Remarks

Presenter: Deputy Chief Data Officer, Marcus Thornton 1:02 PM – 1:08 PM

Discussion: Marcus Thornton welcomed everyone and introduced himself. The purpose of the Data Governance Council is to advise the Chief Data Officer on data technology, policy, and governance structure. The idea is to discuss best practices and standards that can be taken back to the agencies to implement the policies. The council is here to create a standard for agencies to adhere to. In addition, a follow-up from the October Data Governance Council meeting was provided.

Decision: N/A

Agenda Item 2: Roll Call 1:08 PM – 1:09 PM

Presenter: Deputy Chief Data Officer, Marcus Thornton

Discussion: Marcus Thornton called roll.

Decision: The attendee list is presented at the top. Quorum has been met.

Agenda Item 3: Data Governance Council Order of Business 1:09 PM – 1:13 PM

Presenter: Deputy Chief Data Officer, Marcus Thornton

Discussion: The data governance policy and the previous meeting minutes were discussed. No nominations were made for Vice Chair of the council, so that issue has been tabled until the next meeting.

Decision: Two motions were approved. Refer to Motion 1 and Motion 2 presented at the top.

Agenda Item 4: NIEM Overview 1:13 PM – 1:34 PM

Presenter: OASIS-Open, Deputy Division Chief, Katherine Escobar

Discussion: Katherine Escobar provided an overview of NIEMOpen. NIEM is a framework to enable interoperability that was developed in the early 2000's after the 9/11 tragedy. Virginia adopted NIEM in 2012. It is a community-driven, open source, standards-based approach to exchange information. NIEM enables organizations to move information across organizational boundaries to interoperate, and act as one, while each maintains authority for their own existing systems. Outreach, marketing, and trainings for NIEM were discussed.

Key Points covered:

- NIEMOpen Purpose
- Data Interoperability Challenges and Solutions
- NIEMOpen Framework
- Outreach, Training, and Resources

Decision: N/A

Agenda Item 5: Cross-State Report on Data Governance and Literacy 1:34 PM – 2:02 PM
Presenter: The Idea Garden Guest Speaker, Jamie Kimes

Discussion: Jamie Kimes introduced herself as representing the state of Colorado. Research has been conducted by having discussions with twelve different states about their data literacy approach. Among the twelve states, Texas stood out as they have the most advanced data literacy program that is easy to access. Most of the other states utilizes Arizona State University's (ASU) data literacy program for

research. Overall, data literacy can help solve many problems as a lot of businesses ask for services but don't really know what they are asking for.

From a collaborative discussion within the Data Governance Council, it is recommended that any future trainings result in receiving some kind of certification or badge. The trainings should be easy to learn and include gamification. The common three areas all agencies agreed to be beneficial for trainings are:

1. Data Quality
2. Data Stewardship
3. Data for Executive Management

Decision: N/A

BREAK: 2:02 PM – 2:09 PM

Agenda Item 6: Big ID Demo 2:09 PM – 2:20 PM
Presenter: Director of Data Protection and Governance, Chris Burroughs, and Data Protection Analyst, Imran Afridi

Discussion: Big ID is a tool that the ODGA obtained to scan unstructured data. It was piloted with several executive branch agencies which resulted in the discovery of sensitive data. This tool has many different capabilities and can be personalized based on the specific classifiers needed. The unstructured data scanning is such a powerful tool because many agencies don't even know what type of information is found. By identifying files with sensitive information, agencies can better manage the security of their data.

A live demo was presented to the Data Governance Council to demonstrate how the product works and showcase its functionalities.

Decision: N/A

Agenda Item 7: Informatica Data Quality Demo 2:20 PM – 2:36 PM

Presenter: Informatica Solutions Consultant, Rebecca Armstrong

Discussion: Rebecca Armstrong discussed Informatica and the Data Quality Methodology. The main two areas of focus in the state of Virginia are 1) Discover and 2) Measure and Monitor. ODGA is currently negotiating with Informatica for a contract to bring to the state. The hope is to offer this service to the agencies in the Commonwealth.

A live demo was presented to the Data Governance Council to demonstrate how the product works and showcase its functionalities.

Decision: N/A

Agenda Item 8: ODGA Update 2:36 PM – 2:50 PM

Presenter: Deputy Chief Data Officer, Marcus Thornton, Director of Data Protection and Governance, Chris Burroughs, and Acting Director of Communications, Jessi Bailey

Discussion: The Office of Data Governance hosts the Virginia Datathon once a year. Jessi Bailey explained the purpose of the Datathon and gave an overview of this year's theme: "Boosting the Virginia Workforce." Some of the solutions from this event are currently being reviewed and potentially could be implemented to better the Commonwealth.

Marcus Thornton discussed the annual Data Management Maturity Assessment. This is a self-survey given to the agencies that consists of 30 questions to assess their own data maturity. It allows the ODGA to understand their data maturity level and how to engage with each agency, identify gaps, and provide solutions.

The Office of Data Governance and Analytics has multiple projects in place, both internally and externally. Currently the agency is building a Commonwealth-wide data catalog, has improved the Open Data Portal, and increased membership in the Commonwealth Data Trust. In addition, the agency is working on multiple projects, including other agencies, such as the Department of Forestry (DOF), the Department of Forensic Science (DFS), the Virginia Community Engagement Index (VCEI), and the Virginia Permit Transparency (VPT). These projects have resulted in increased efficiency through automation and significant monetary savings.

Key Points covered:

- 2024 Virginia Datathon
- Data Management Maturity Assessment
- ODGA Data Catalog
- New and Improved Open Data Portal
- ODGA Projects

Decision: N/A

Agenda Item 9: Member and Public Comment 2:50 PM – 2:51 PM

Presenter: Deputy Chief Data Officer, Marcus Thornton

Discussion: Marcus Thornton opened the floor to questions and comments. No additional comments were made.

Decision: N/A

Agenda Item 10: Closing Remarks 2:51 PM

Presenter: Deputy Chief Data Officer, Marcus Thornton

Discussion: Marcus Thornton thanked everyone for their attendance and participation.

Decision: N/A

Agenda Item 11: Adjourn 2:51 PM

Presenter: Deputy Chief Data Officer, Marcus Thornton

Discussion: The Data Governance Council meeting adjourned at 2:51 PM.

Decision: N/A

COV Data Governance Guidelines

Purpose

The purpose of this policy is to establish a framework for the management and protection of data within executive branch agencies, ensuring data quality, security, and accountability.

Scope

This policy applies to all executive branch agencies and their employees who handle, manage, or access data.

Governance Committees

Data is managed by a multi-tiered hierarchy

Committee	Purpose
Virginia Data Advisory Commission	The Virginia Data Advisory Commission (the Commission) is established as an advisory commission in the executive branch of state government to advise the Office of Data Governance and Analytics (the Office) on issues related to data sharing, including open data, data analytics, and data governance. The Commission shall (i) set, plan, and prioritize data sharing performance goals for the Commonwealth, (ii) review agency accomplishments, and (iii) recommend solutions that will establish the Commonwealth as a national leader in data-driven policy, evidence-based decision making, and outcome-based performance management. Virginia Data Advisory Commission Charter
Executive Data Board	The Executive Data Board is chaired by the Chief Data Officer who also selects its membership. The Executive Data Board membership consists of executive leadership, or their designees, from executive branch agencies engaged in data sharing and analytics projects. The Board has the following responsibilities: • Translate the Commonwealth's data-driven policy goals and objectives to agency performance targets • Allocate appropriate agency resources to support data governance, sharing, and analytics initiatives

Committee	Purpose
	• Provide to the Virginia Data Commission any reports on the Board's recommendations and work as required by the Commission
Data Governance Council	The members of the Data Governance Council are selected by the Executive Data Board. The Council, however, consists of employees of the state agencies represented on the Executive Data Board. The Data Governance Council is chaired by the Commonwealth of Virginia Chief Data Officer (or designee) and has the following responsibilities: • Liaise between state agency operations and the Chief Data Officer • Advise the Chief Data Officer on technology, policy, and governance strategies • Administer data governance policies, standards, and best practices as set by the Executive Data Board • Oversee data sharing and analytics projects • Review open data assets • Govern the Commonwealth Data Trust • Report progress, compliance, and performance to the Executive Data Board
Data Stewards Group	The objectives of the Data Stewards Group are to: • Promote and facilitate the secure and appropriate sharing and use of data assets in support of data-driven policymaking, evidence-based decision-making, research, and analysis • Maximize the value and utility of Commonwealth data-related investments and assets • Promote increased sharing of data between state agencies and localities to provide tangible operational improvements to assist state agencies and localities in fulfilling their respective missions in a coordinated, cost-effective manner • Provide public access to data assets, where lawful and appropriate, to enhance research, innovation, and insight

Note: The Data Sharing and Analytics Advisory Committee has been disbanded.

Roles and Responsibilities

Per COV (§2.2-603.F) "the director of every agency and department in the executive branch of state government, including those appointed by their respective boards or the Board of Education, shall be responsible for securing the electronic data held by his agency or department and shall comply with the requirements of the Commonwealth's information technology security and risk-management program as set forth in § 2.2-2009." In addition, the Director of every department is responsible for the security of the agency's electronic information, and for establishing and maintaining an agency

information security program compliant with this policy and meets all of the requirements established by COV ITRM Security Standards.

Policy

Role Definition

- Agencies have clearly defined key roles of Data Owner, Data Steward, and Data Custodian.

Key Data Management Roles

Role	Definition
Chief Data Officer (CDO)	• The Chief Data Officer is responsible for the overall management of the COV's Data and Information Governance
Agency Senior Management	• Group in charge of deciding whether to publish the agency's data assets. A Data Executive supported by a Data Owner has the responsibility for the management of data assigned within their portfolio.
Data Owner	• The data owner in an organization is anyone who originally created or acquired the raw data, and/or anyone who retains the rights to approve sharing or other actions, can make access control decisions (or restrictions), and can approve memorandums of agreement on data use.
Data Steward	• Data stewards carry out data owner requirements, translate them into meaningful requirements, and act as subject matter experts on the data and its utility for business use. Where data owners have decision rights, stewards protect those rights and convey them across the enterprise. Data stewards bridge the gap between different data stakeholders, so they need excellent communication skills in order to translate between data owners, platform developers, and data users.
Data Custodian	• A data custodian is a particular type of data steward. In some organizations, a data custodian is referred to as a technical data steward. The role of the data custodian is closer to the IT side of the house than the mission side, and typically covers tasks such as information flow, operation of APIs, and articulating technical data requirements to platform development teams.

•

Data Catalog

- Each agency must maintain a comprehensive data inventory that lists all data assets under its control, including data sets, databases, files, and data sources.
- The data catalog must list all data assets with descriptions, data classifications, and responsible data stewards.
- The data catalog must be regularly updated and include metadata such as data source, data owner, data classification, and data retention policies.
- The data catalog should be accessible to authorized personnel within Commonwealth agencies for reference and discovery of available data resources.
- Datasets submitted to Archer can be published as the agency's data catalog or agencies may request ODGA to scan their assets into the Commonwealth of Virginia's Data Catalog.

Data Classification

- All datasets must be classified based on its level of sensitivity and criticality of the data
- Data classification labels should include but are not limited to Public, Internal Use Only, Confidential, and Highly Sensitive
- Data shared with ODGA must identify a Tier level (see table below) to ensure the proper security and protection of data.
- Data owners are accountable for ensuring the proper classification of data and data stewards are responsible for communicating the Tier level of data provided to ODGA.

Tier	Description
Tier 1 Data	Data that is not protected from public disclosure or subject to withholding under any law, regulation, or contract. Nevertheless, publication of the dataset on the public Internet and exposure to search engines would: have the potential to jeopardize the safety, privacy, or security of a person who may be identified through use of the data; requires subjective redaction to classify the data as Tier 0 data; impose an undue financial or administrative burden on the Data Trust Member; or expose the Trustee or Data Trust Member to litigation or liability.
Tier 2 Data	Sensitive or proprietary information intended for access or release only on a 'need-to-know' basis, including personal information not otherwise classified as Tier 0 or 1, and data protected or restricted by contract, grant, or other agreement terms and conditions provided on an ongoing basis or as a one-time transfer to Trustee by Data Trust Member for use by Data Trust under this Agreement as detailed in Exhibit B attached hereto or subsequently contributed by Data Trust Member and detailed in the Data Trust electronic metadata registry.
Tier 3 Data	Sensitive or proprietary information and data elements with a statutory requirement under Data Trust Member's relevant state and federal laws for notification to affected parties in case of a confidentiality breach (e.g. Social Security Number, driver's license number, financial account numbers,

Tier	Description
	personal medical information, etc.) provided on an ongoing basis or as a one-time transfer to Trustee by Data Trust Member for use by Data Trust under this Agreement as detailed in Exhibit B attached hereto or subsequently contributed by Data Trust Member and detailed in the Data Trust electronic metadata registry. Examples of Tier 3 Data may include, but not limited to: Attorney-Client Privileged; Criminal Justice Information; Critical Infrastructure Information; Family Educational Rights and Privacy Act (FERPA); Federal or State Tax Information; or Protected Health Information (PHI) under the Health Insurance Portability and Accountability Act (HIPAA).
Tier 4 Data	Sensitive or proprietary data where the unauthorized disclosure could potentially cause major damage or injury, including death, to entities or individuals identified in the information, or otherwise significantly impair the ability of the Data Trust Member to perform its statutory functions. Tier 4 Data includes any dataset designated by a federal agency at the level "Confidential" or higher under the federal government's system for marking classified information. No Tier 4 data shall be knowingly incorporated into the Data Trust.

Key Data Governance Policies

- Each agency must establish and adhere to a set of key data governance polices including, but not limited to:

Policy	Minimum Requirements
Data Stewardship	• Each data asset must have an assigned data steward responsible for its quality, security and compliance. • Data stewards must ensure data is used for its intended purposes and enforce data access and usage policies.
Data Quality	• Agencies must establish data quality standards and regularly assess and improve data quality. • Data quality assessments should include data accuracy, completeness, consistency, and timeliness.
Data Security	• Data security measures must be in place to protect data assets from unauthorized access, disclosures, alteration, and destruction. • Encryption, access controls, and data masking should be implemented based on data classification.
Data Retention	• Agencies must develop data retention policies to ensure data is retained for an appropriate period. • Data retention schedules must comply with legal and regulatory requirements.
Data Privacy	• Agencies must adhere to data privacy laws and regulations, protecting the privacy of individuals' data. • Personal data must be handled with care and in compliance with relevant privacy standards.

Reporting and Metrics

- Agencies must maintain a data risk register of data management issues,
- Agencies must establish a system for collecting data governance metrics.
- Metrics should include, but are not limited to, data quality, data risks, and compliance with data governance policies.

Sample Metrics

- **Data Quality**
  - Data accuracy – The percentage of data records that are error-free and consistent with the source
  - Data Completeness – The percentage of expected data elements present in a dataset

        - Data Consistency – The level of uniformity in data values and formats across different systems and sources
        - Data Timeliness – The average time it takes to update and make data available for use.
    - **Data Security**
        - Data access controls – The number of unauthorized access attempts or breaches
        - Data encryption – The percentage of sensitive data that is encrypted at rest and in transit
        - Data masking effectiveness – The number of successful attempts to access masked or redacted data
        - Incident response time – The average time it takes to detect and respond toa data security incident
    - **Data Privacy**
        - Data privacy compliance – The percentage of data assets that comply with relevant data privacy regulations (e.g. HIPAA)
        - Data privacy training – The percentage of agency staff who have completed data privacy training
    - **Data catalog**
        - Data Inventory completeness – The percentage of data assets included in the data catalog
        - Data catalog usage – The number of searches, views, and downloads of data catalog entries
        - Data catalog updates – The frequency and timeliness of data catalog updates
    - **Data Stewardship**
        - Data steward assignments – The percentage of data assets with assigned data stewards
        - Stewardship activities – The number of data stewardship activities, including data quality assessments, data audits, and policy enforcement
        - Data steward training – The percentage of data stewards who have completed relevant training
- Data Governance Status Reports with key metrics must be submitted to the Commonwealth of Virginia Chief Data Officer, at least annually.
- These reports should include recent successes, areas of improvement and actions taken to address data governance issues.

Policy Review

This Policy will be reviewed and updated every three (3) years from the approval date, or more frequently if appropriate. Any staff members who wish to make any comments about the Policy may forward their suggestions to the Director of Data Governance, ODGA.

Related Policies

ODGA Policies, Standards and Procedures
Data Protection Policy
Metadata Policy

The Office of Data Governance and Analytics adheres to all Commonwealth Information Technology Resource Management (ITRM) policies and standards for security and architecture Policies, Standards & Guidelines | Virginia IT Agency.

VITA Related Policies
IT Information Security Policy - SEC519
Information Security Standard (SEC501)
IT Risk Management Standard (SEC520)

Version History

Version Number	Revision Date	Description of Change	Author
V1	10/31/2023	Initial Draft	Chris Burroughs

Data Governance Council

May 2024

Agenda

1:00 p.m.- Welcome and Opening Remarks
Marcus Thornton, Deputy Chief Data Officer

1:05 p.m.- Roll Call

1:10 p.m.- DGC Order of Business
- Approving previous meeting minutes
- Election of Vice Chairperson
- Approval of Data Governance Policy
- Member Status- Purview and ODP

1:15 p.m.- NIEM Overview

1:30 p.m.- Cross State Report on Data Governance and Literacy

1:45 p.m.- Big ID Demo

2:00 p.m.- Break

2:05 p.m.-ODGA Update

2:20 p.m.- Informatica Data Quality Demo

2:35 p.m.- Member and Public Comment

2:45 p.m.- Adjourn

odga

Data Governance Council

Purpose

Advise the CDO on data technology, policy, and governance structure.

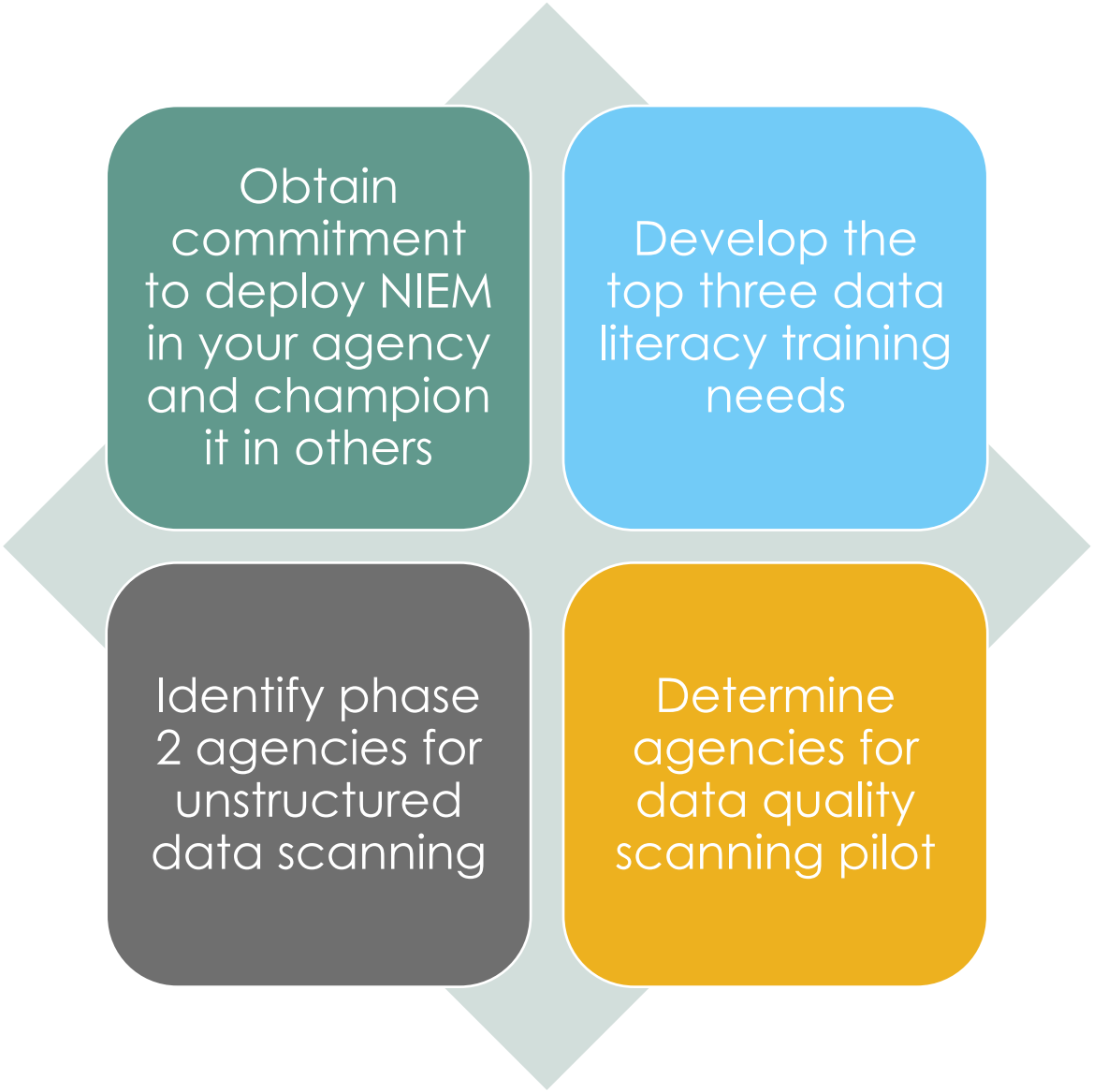
Administer data governance policies, standards, and best practices, as set by the Board.

Oversee data sharing and analytics projects.

- Liaise between state agency operations and the CDO
- Review open data assets prior to publication.
- Provide to the Board any reports on the Council's recommendations and work as required by the Board.
- Develop necessary privacy and ethical standards and policies for Commonwealth Data Trust resources.
- Monitor the sharing of Commonwealth Data Trust member-contributed data resources.
- Review and approve new Commonwealth Data Trust-managed data resources.
- Conduct any other business the CDO deems necessary for Commonwealth Data Trust governance.

Related legislation: https://lis.virginia.gov/cgi-bin/legp604.exe?212+ful+CHAP0314

Meeting Purpose



Order of Business

1. Approve Meeting Minutes
2. Select Vice Chairperson
3. Approve Data Governance Policy

Follow-up from Oct 2023 Meeting

- We appreciate the agencies that updated Purview:
  - VDH
  - DWR
  - VEC

- Update Purview data owners and data stewards
- Review Purview metadata and rate data sets with stars
- Review Open Data Portal data and increase the number of datasets for your agency

Open Data Portal Contributions

Agency	# Datasets
Virginia Department of Environment Quality	84
City of Norfolk	78
Virginia Department of Health	74
Department of Housing & Community Development	66
Virginia Commonwealth University	51
City of Dumfries	39
Library of Virginia	23
Office of Intermodal Planning and Investment	23
Department of Elections	10
Department of General Services	9
Department of State Police	7
Virginia Housing	2
Department of Conservation and Recreation	1

NIEM Overview

Katherine Escobar

odga

5/31/2024

AGENDA

- Why NIEMOpen
- What Is NIEMOpen
- NIEMOpen Project
- Questions

17 NIEMOpen Domain Spaces represented by NBAC TSC Sub-Committees

WHY NIEMOPEN?

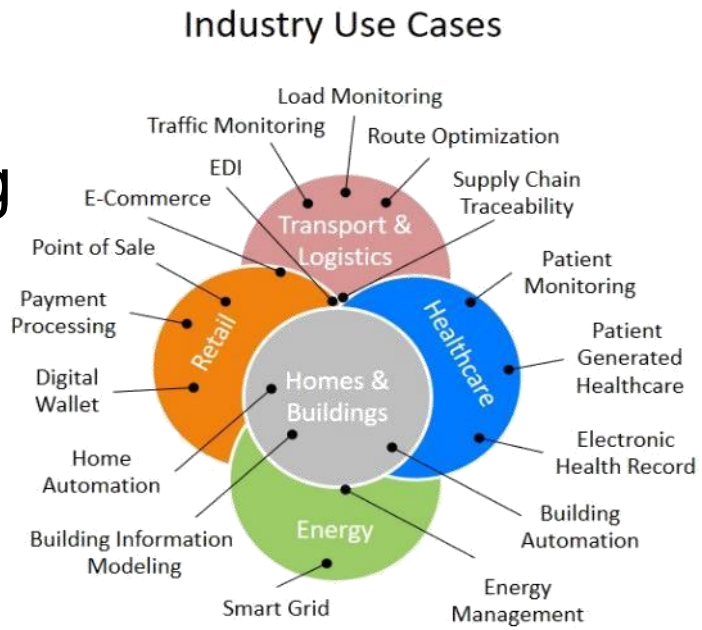
- Our mission is furthered by the sharing of data:
  - With Federal, State, Local, Tribal, Territorial, International partners (ex: DHS, FBI, etc.)
  - With industry and vendor partners
  - With compliance drivers (ex: HL7, LOINC, SNOMED for medical, ASC X12, etc.)

- For Interagency, this includes:
  - Multiple, complex data sharing permutations via forms, digital, emails, phone calls, Slack, and other mechanisms
  - Data must be shared across multiple networks, multiple security boundaries, multiple IT capabilities, and in multiple ways, but still be understandable
  - Component programs are not equipped to handle all domains in this Enterprise-level data sharing challenge

DATA INTEROPERABILITY

Across all levels of government, industries, and technical solution providers, **interoperability** is an essential tool for:

- Better decision making
- Reduction of manual processing
- Increased productivity
- Reduction of errors
- Management of costs



Everyone agrees it takes interoperability of systems to achieve these benefits.

NIEMOPEN

COMMON INTEROPERABILITY CHALLENGES

If everyone agrees, why is there a lack of interoperability problem?

- Lack of unified agreement on standards selection and implementation
- Siloed (domain-specific) development efforts driven by need, isolated by incompatibility
- Proprietary and vendor-driven solutions vs. Interoperable driven solutions
- Information Blocking - business, technical, and organizational practices that prevent or materially discourage access, exchange or use

INTEROPERABILITY SOLUTION

The key factor to achieve interoperability is **data**

- Information exchange is critical to ensuring timely information is available when and where it is needed
- To exchange data, systems must have agreement on:
  - **Syntactic**: Data encodings and representation
  - **Semantic**: Consistent terminology and meanings

NIEMOpen Solves the Interoperability Challenge at the Data Level

NIEMOPEN

DATA INTEROPERABILITY PROBLEM

Cyber Terms
SPAM
Honeypot
Worm

DATA UNDERSTANDING (Semantics, Syntax, Structure)

IF THESE AREN'T COMPATIBLE…

THIS WON'T WORK

DATA UNDERSTANDING (Semantics, Syntax, Structure)

SPAM ≠ 

Honeypot ≠ 

Worm ≠ 

SYSTEM / APPLICATION

DATA EXCHANGE @ RUNTIME

SYSTEM / APPLICATION

Data Producers

Data Consumers 

NIEMOPEN

NIEMOPEN

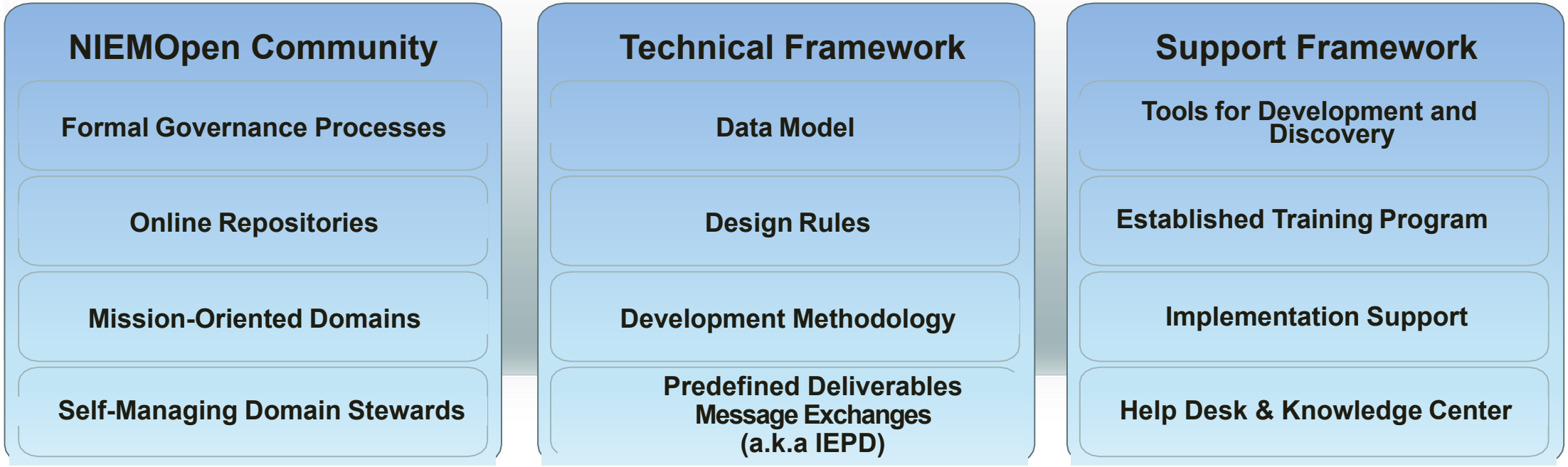
What is NIEMOPEN

21 May 2024

Ms. Katherine Escobar
NIEMOpen Project Governing Board Chair

NIEMOPEN FRAMEWORK

- NIEMOpen is a community-driven, open source, standards-based approach to exchanging information
- Diverse communities can collectively leverage NIEMOpen to increase efficiencies and improve decision-making
- NIEMOpen is available to everyone, including public and private organizations
- NIEMOpen includes a data model, governance, training, tools, technical support services, and an active community to assist users in adopting a standards-based approach to exchanging data



NIEMOPEN

NIEM – THE BIG PICTURE

NIEM enables organizations to move information across organizational boundaries to interoperate – and act as one – while each maintains authority for their own existing systems

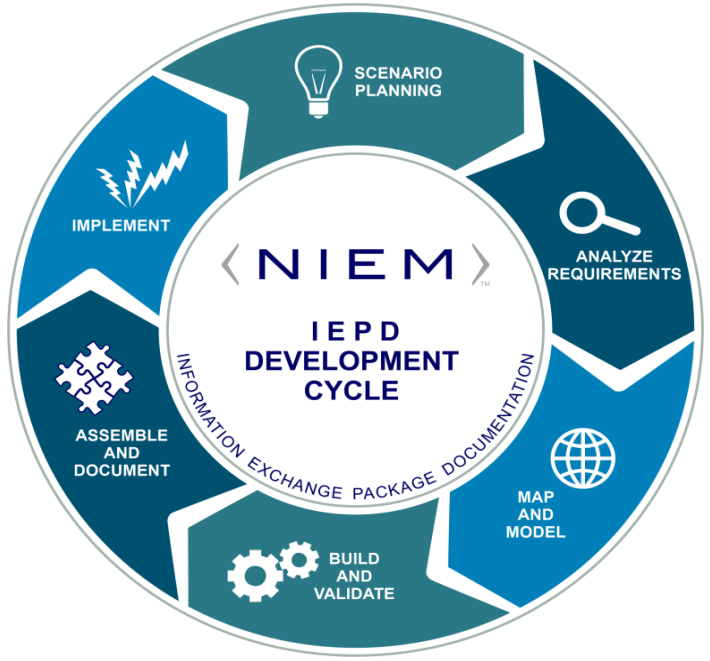
Common Language

(Data Model Lifecycle)

Built and governed by users within the Federal, State, Local, Tribal, and Private Sectors

Repeatable, Reusable Process

(Exchange Specification Lifecycle)



NIEMOPEN

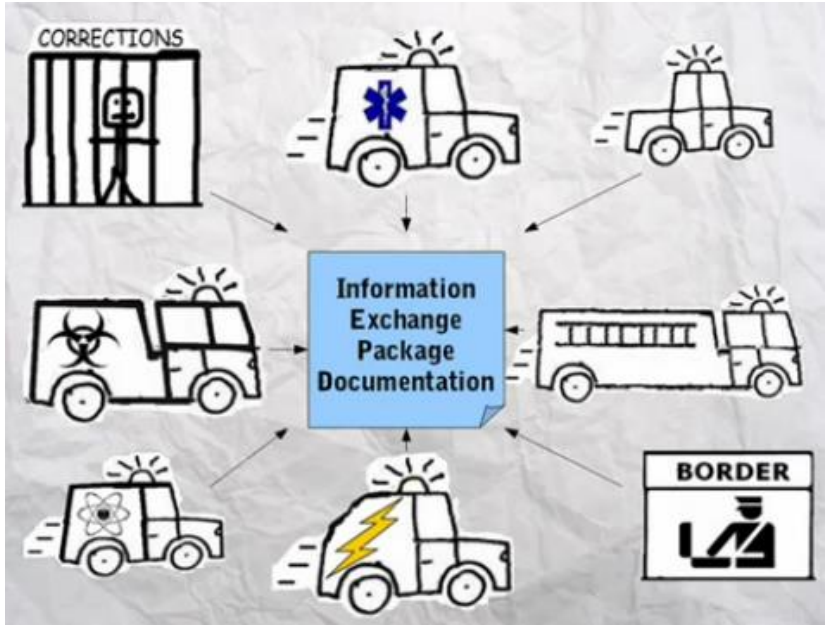
NIEMOPEN ADVANTAGE

When using NIEM, you only need to "speak" two languages — your own and NIEM

EXAMPLE Communicating With 15 Different Systems $N^{(2)}$ Problem

WITHOUT NIEM

USING NIEM



NIEMOPEN

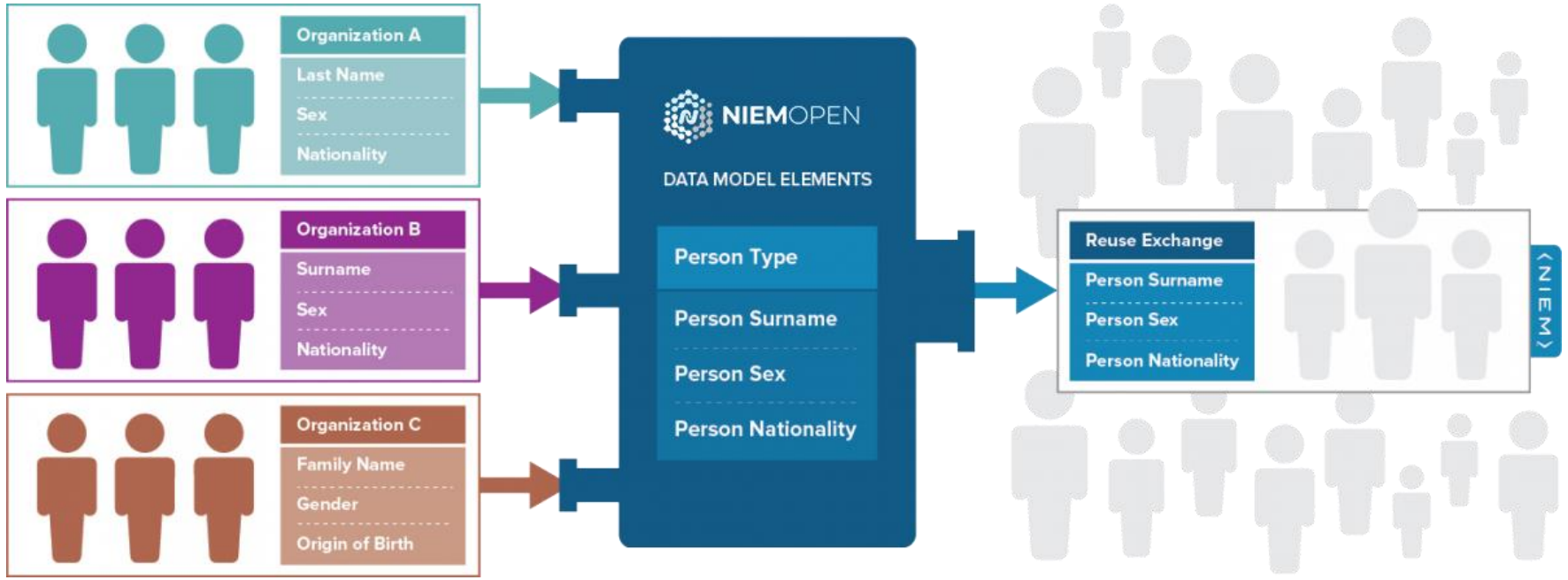
NIEMOPEN IN ACTION:
How NIEMOpen Exchanges Data Components and Data Elements

NIEMOpen is

a common vocabulary –
enabling efficient information exchange across diverse public and private organizations

NIEMOpen is not

a system or database –
NIEMOpen does not specify how to transmit or store data



Using NIEM, organizations come together to agree on a common vocabulary
When additional organizations are added to the information exchange, the initial NIEMOpen exchange can be reused, saving time and money

NIEMOPEN

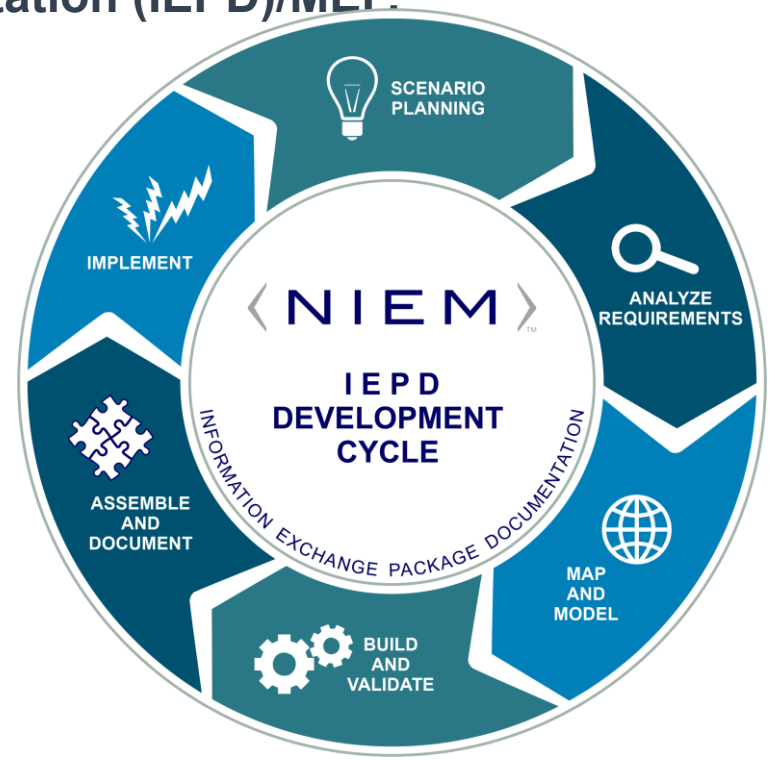
WHAT IS A NIEM INFORMATION EXCHANGE?

In NIEM, an information exchange is also known as an **Information Exchange Package (IEP)/Message Exchange Package (MEP),** a description of specific information exchanged between a sender and a receiver.

The IEP/MEP is usually coupled with additional documentation, sample message instances, business rules, and more to compose an **Information Exchange Package Documentation (IEPD)/MEP.**

Core Functions of an IEPD/MEP

- Developed to provide the business, functional, and technical details of the information exchange through predefined artifacts
- Created with a core set of artifacts in a prescribed format and organizational structure to allow for consistency
- Designed to be shared and reused in the development of new information exchanges through the publication in IEPD repositories



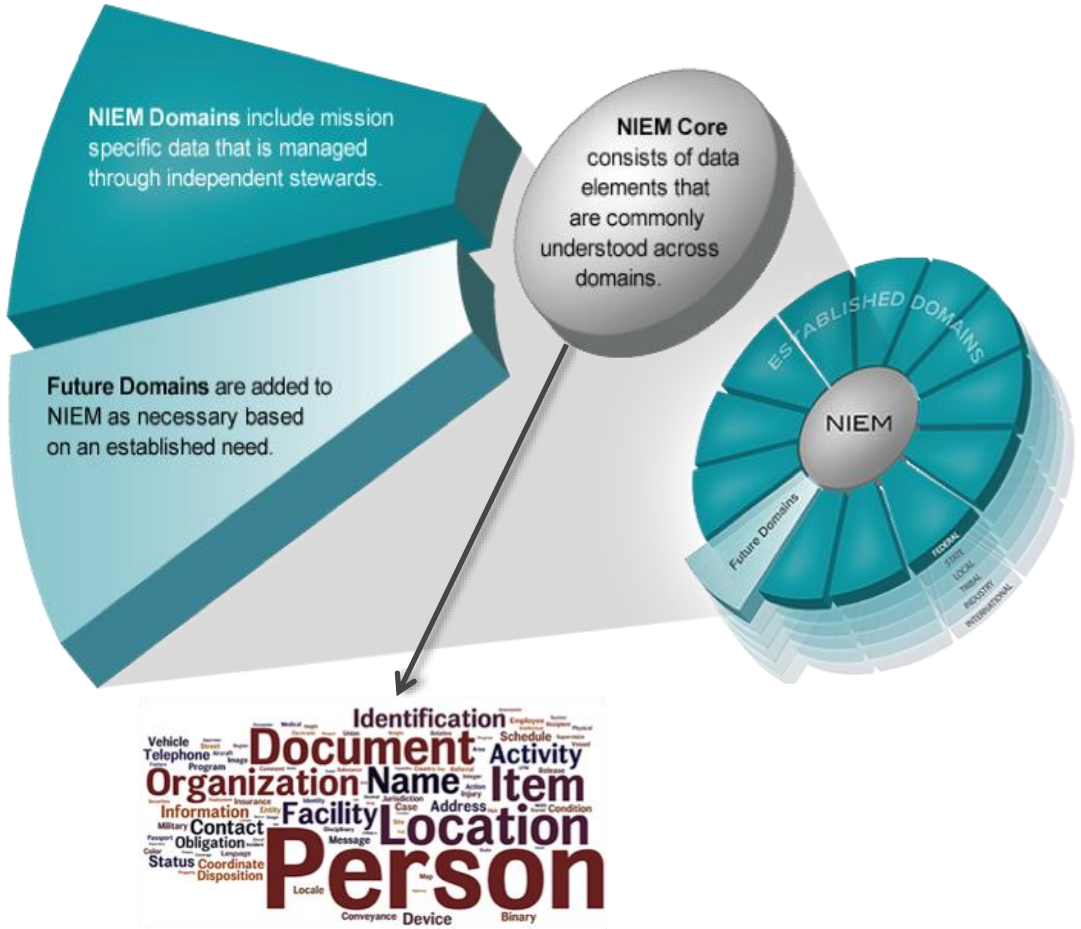
NIEMOPEN

NIEMOPEN MODEL CONTENT

Think of the NIEMOpen data model as a mature and stable data dictionary of agreed-upon terms, definitions, and formats independent of how information is stored inside individual IT systems.

NIEMOpen 17 Domains

- Agriculture
- Biometrics
- Chemical, Biological, Radiological, & Nuclear
- Cyber
- Emergency Management
- Human Services
- Immigration
- Infrastructure Protection
- Intelligence
- International Trade
- International Human Services
- Learning & Development
- Justice
- Maritime
- Military Operations
- Screening
- Surface Transportation

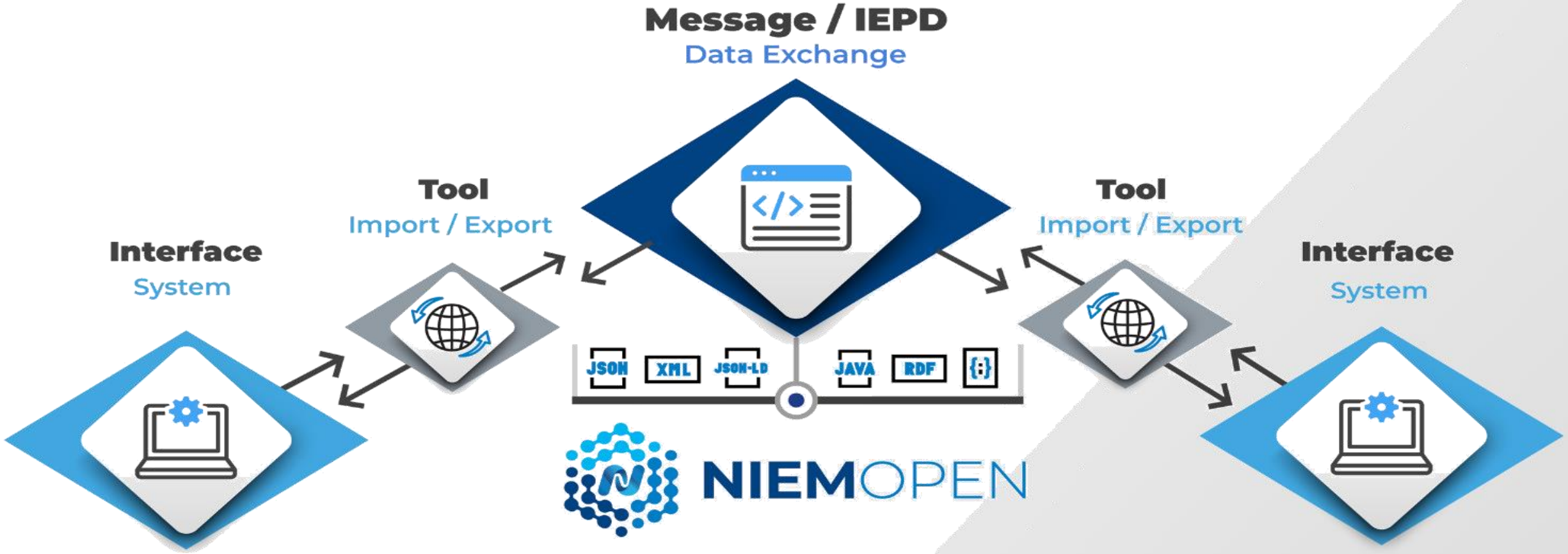


Emerging Domains

- Acquisition – DoD A & S
- Identity Credentials – DHS
- Physical Security - DMDC

NIEMOPEN

Scope of NIEM



NIEM is a data layer standard and intentionally does not address all the necessary technologies needed for information sharing.

Exchange partners decide how to store and process the NIEM-conformant data being exchanged.

NIEMOPEN

NIEMOPEN

NIEMOPEN Project

21 May 2024

Ms. Katherine Escobar
NIEMOpen Project Governing Board Chair

NIEMOpen Current SPONSORS “No Change”

9 Sponsors, 3 Premier

- **Sponsors**
  - JS J6
  - DHS S&T
  - CJIS (FBI)
  - GTRI
  - NAJIS
  - IJIS
  - DOT
  - **ODGA (Virginia)**
  - Equivant

Premier

Premier

Premier

- **In Discussions**
  - Microsoft
  - DHS OBIM
  - DND, Canada
  - DOJ

- **Expressed Interest**
  - MITRE
  - Veterans Administration
  - Tyler technologies
  - Ernst & Young
  - NIST
  - DISA
  - NOAA
  - Pinary

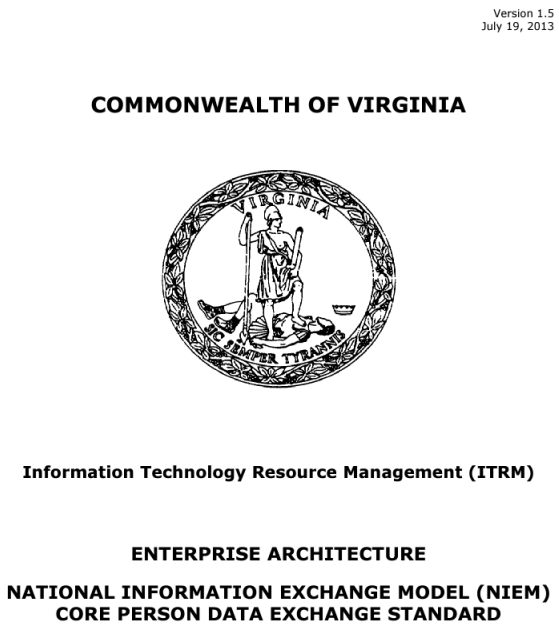
Department of National Defense

Ms. Carol Geyer

– OASIS Open Chief Development Officer

NIEMOPEN

VIRGINIA ADOPTION OF NIEM


Version 1.5
July 19, 2013

COMMONWEALTH OF VIRGINIA

Information Technology Resource Management (ITRM)

ENTERPRISE ARCHITECTURE

NATIONAL INFORMATION EXCHANGE MODEL (NIEM)
CORE PERSON DATA EXCHANGE STANDARD

Commonwealth of Virginia's EIA Strategy and NIEM Integration Plan:
Commonwealth of Virginia has completed an eight-month strategic planning process to develop an Enterprise Information Architecture (EIA) strategy. A central element of the EIA strategy involves building exchanges for "citizen-centric" data that conform with the National Information Exchange Model (NIEM).

The following section references the statutory authority granted by the *Code of Virginia* to the Secretary of Technology to establish the NIEM Core Person Data Exchange Standard. The section also cites provisions in the *Code of Virginia* relating to the role of the Chief Information Officer (CIO) of the Commonwealth, the Information Technology Advisory Council (ITAC) and HITSAC in the development, review and adoption of Commonwealth ITRM Standards.

The Secretary of Technology
§ 2.2-225. Position established; agencies for which responsible; additional powers
http://leg1.state.va.us/cgi-bin/legp504.exe?000+cod+2.2-225

The Chief Information Officer (CIO) of the Commonwealth
§ 2.2-2007. Powers of the CIO
http://lis.virginia.gov/cgi-bin/legp604.exe?000+cod+2.2-2007

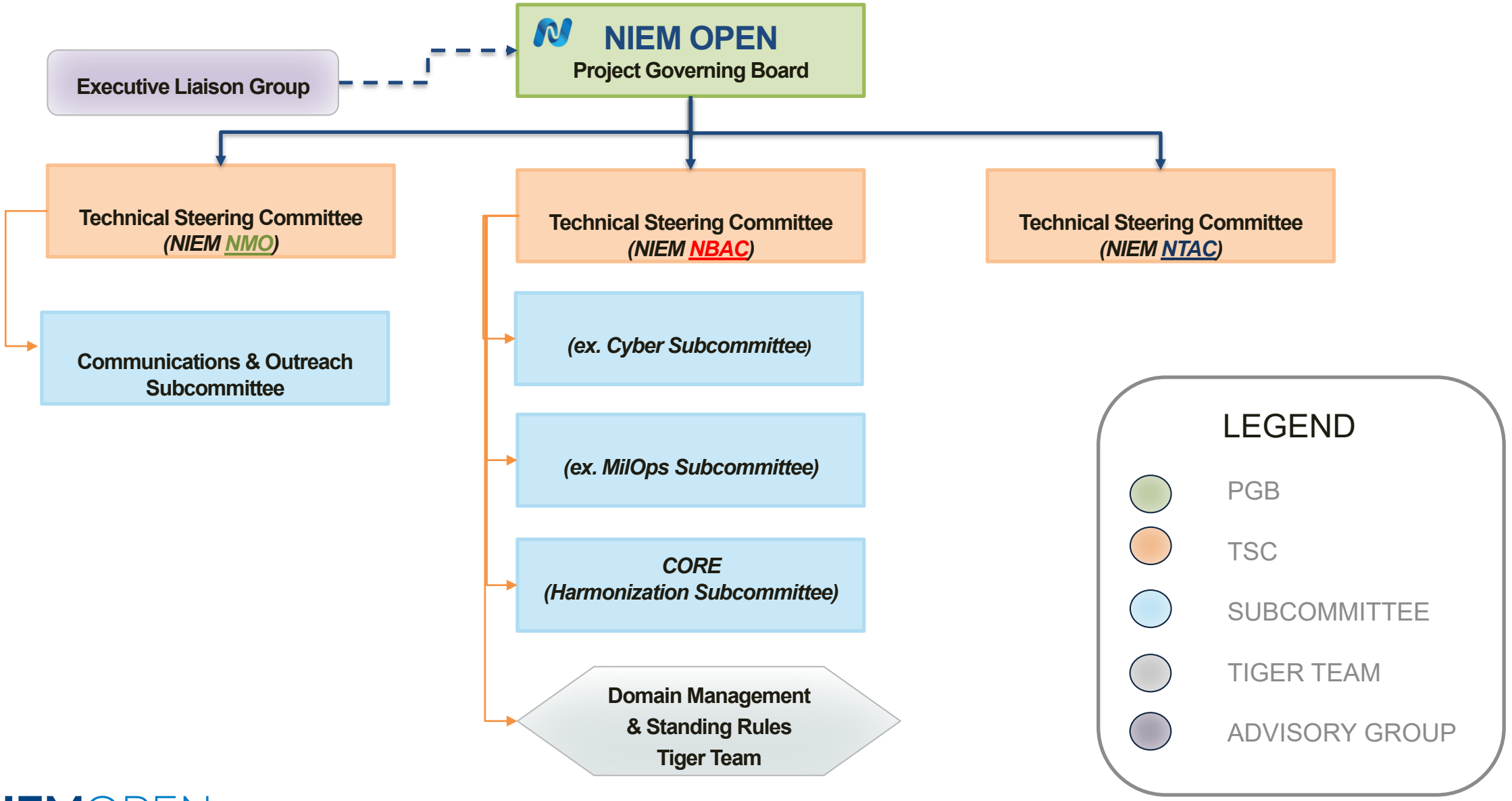
The Virginia Information Technologies Agency
§ 2.2-2010. Additional powers of VITA
http://lis.virginia.gov/cgi-bin/legp604.exe?000+cod+2.2-2010

The Information Technology Advisory Council (ITAC)
§ 2.2-2699.6. Powers and duties of the ITAC
http://lis.virginia.gov/cgi-bin/legp604.exe?000+cod+2.2-2699.6

The Health Information Technology Standards Advisory Committee (HITSAC)
§ 2.2-2699.7. Health Information Technology Standards Advisory Committee
http://leg1.state.va.us/cgi-bin/legp504.exe?000+cod+2.2-2699.7

NIEMOPEN

NIEM OPEN PROJECT GOVERNANCE



PGB APPOINTED VOTING MEMBERS

Ms. Katherine Escobar
PGB Chair
Deputy Division Chief, Data and Standards Division at Joint Staff J6 and NIEM Managing Director

Ms. Christina Bapst-Stump
Science and Technology Directorate, US Department of Homeland Security

Mr. Gary Egner
Director Business Development, Equivant

Mr. John Wandelt
Principal Research Scientist, Georgia Tech Research Institute

Ms. Maria Cardiellos
Executive Director Integrated Justice Information Systems Institute

Mr. Payton Lamb
Data Engineer Commonwealth of Virginia
Office of the secretary of Administration
Office of Data Governance and Analytics

Ms. April Mitchell
FBI It Specialist and Technical Lead of Data Standards in the Criminal Justice Information Services (CJIS) Division

PGB EXPERT VOTING MEMBERS

Technical Steering Committees

Kamran Atri
Co-Chair
NIEM Business Architecture Committee (NBAC) Technical Steering Committee (TSC)

Jim Cabral
Vice President, Court Relations, InfoTrack

Beth Smalley
Co-Chair
NIEM Management Office (NMO) Technical Steering Committee (TSC)

PGB NON-VOTING MEMBERS

Technical Steering Committees

Thomas Krul
Co-Chair
NIEM Business Architecture Committee (NBAC) Technical Steering Committee (TSC)

Dr. Scott Renner
Co-Chair
NIEM Technical Architecture Committee (NTAC) Technical Steering Committee (TSC)

NIEMOPEN

OUTREACH/TRAINING

- Marketed / Booked Training Classes (500+ attendees)

Total Tickets	Category																			
YR-QTR	Academia	Canada	City/State /Local	DHS	DoD	DoJ	EIA	First NEt	MIT RE	Oth er	Othe r-.org	DOC	NASA	DOE	DOT	USDA	Finland	IRS	India	ODNI	Grand Total
2023-Q4	1	1	1		1					1					5						10
2023-Q3			2	1	4					9	1										17
2023-Q2	1		2	1	18	1				14											37
2023-Q1	1	5	7	4					1	27	4	1	2								52
2022-Q4	2		10	5	3				1	34	1			1	9	1					67
2022-Q3	2		10	4	2	1				42	7				3		1				72
2022-Q2	1	2	32	7	26	4			1	61	3				1			1	1	1	141
2022-Q1	2	17	11	2	16	8	1	1	9	45	3	2									117
2021-Q4		22	1	1	3					3	2										32
2021-Q3		8			17	1			3	3	2										34
Grand Total	10	55	76	25	90	15	1	1	15	239	23	3	2	1	18	1	1	1	1	1	579

- NIEMOpen.org Content Updates
- MEP Builder AWS Hosting Deployment
- EIN Presswire Publication Release
- NIEMOpen.org Analytics integration

Upcoming:

- Marketing of NIEM MEP Builder v3.2.0
- Migrating content from NIEM.gov ➔ NIEMOpen.org; slim down NIEM.gov to focus on gov stakeholders
- Training for NIEM 6.0 release
- NIEM Summit 2025 logistics
- New Sub-committee page layouts; collaboration features
- LearnPress LMS roll out

NIEMOpen Resource Statistics (CY23)

YouTube

37,794
impressions

SOURCEFORGE

417
MEP Builder downloads

NIEMOPEN
https://niemopen.org/

19,294
views

NIEMOPEN

RESOURCES

Katherine Escobar, katherine.b.escobar.civ@mail.mil, 757-203-8631

- **Git Repos**
  - NIEMOpen: https://github.com/niemopen/
  - NTAC TSC
    - https://github.com/niemopen/ntac-admin
  - NBAC TSC
    - https://github.com/niemopen/nbac-admin
  - NMO TSC
    - https://github.com/niemopen/nmo-admin
    -
  - NIEMOpen Slack Channel: https://niemopen.slack.com/
- **Mailing Lists**
  - NIEMOpen: https://lists.oasis-open-projects.org/g/niemopen
  - PGB: https://lists.oasis-open-projects.org/g/niemopen-pgb
  - NTAC TSC: https://lists.oasis-open-projects.org/g/niemopen-ntactsc
  - NBAC TSC: https://lists.oasis-open-projects.org/g/niemopen-nbactsc
  - NMO TSC: https://lists.oasis-open-projects.org/g/niemopen-nmotsc

OASIS OPEN

NIEMOPEN

https://www.niem.gov/

https://niemopen.org/

GitHub

slack

Confluence

Follow Us On Twitter

https://twitter.com/NIEMconnects?ref_src=twsrc%5Etfw

https://www.linkedin.com/groups/1903175/profile

https://www.youtube.com/channel/UCg9qV22PXLBjG41hc-EwVrQ

Message Exchange Package (MEP Registry & Repository

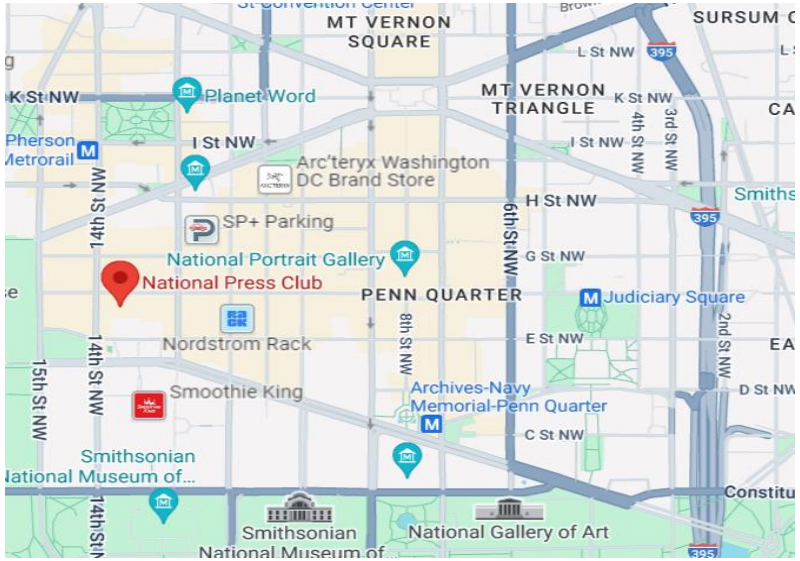
NIEM Message Exchange Package Builder

NIEMOPEN

NIEMOPEN REVEAL SUMMIT : 18-20 FEB 2025

Event Website: https://niemopen.org/reveal

Direct Registration Link: https://www.eventbrite.com/e/niemopen-reveal-training-summit-2025-tickets-890279729637



Homepage | National Press Club

Located in: National Press Building
Address: 529 14th St NW, Washington, DC 20045
Founded: March 29, 1908
Phone: (202) 662-7500

THE NATIONAL PRESS CLUB

NIEMOPEN

NIEMOPEN

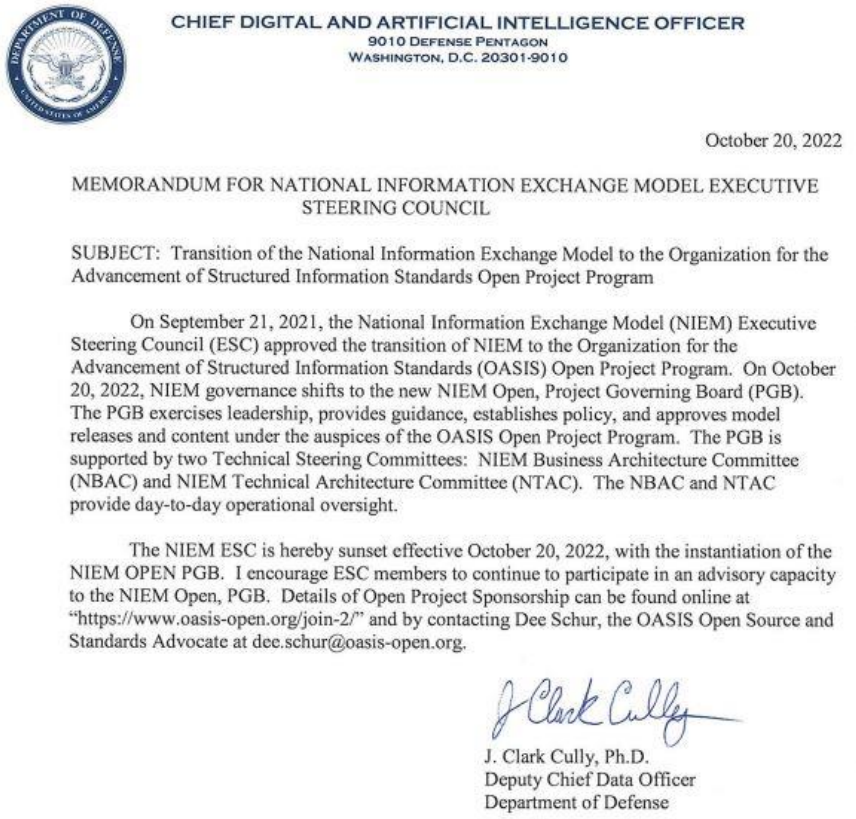
???? Questions ????

21 May 2024

Ms. Katherine Escobar
NIEMOpen Project Governing Board Chair

TRANSFER OF GOVERNANCE

- 20 October 2022 NIEM ESC is Sunset and NIEM Governance Transitions to the NIEM OPEN Project Governing Board (PGB) under the Auspices of OASIS OPEN Project
- ESC Principals remain in an Advisory Role
- Dr. Cully Transition Memo



CHIEF DIGITAL AND ARTIFICIAL INTELLIGENCE OFFICER
9010 DEFENSE PENTAGON
WASHINGTON, D.C. 20301-9010

October 20, 2022

MEMORANDUM FOR NATIONAL INFORMATION EXCHANGE MODEL EXECUTIVE STEERING COUNCIL

SUBJECT: Transition of the National Information Exchange Model to the Organization for the Advancement of Structured Information Standards Open Project Program

On September 21, 2021, the National Information Exchange Model (NIEM) Executive Steering Council (ESC) approved the transition of NIEM to the Organization for the Advancement of Structured Information Standards (OASIS) Open Project Program. On October 20, 2022, NIEM governance shifts to the new NIEM Open, Project Governing Board (PGB). The PGB exercises leadership, provides guidance, establishes policy, and approves model releases and content under the auspices of the OASIS Open Project Program. The PGB is supported by two Technical Steering Committees: NIEM Business Architecture Committee (NBAC) and NIEM Technical Architecture Committee (NTAC). The NBAC and NTAC provide day-to-day operational oversight.

The NIEM ESC is hereby sunset effective October 20, 2022, with the instantiation of the NIEM OPEN PGB. I encourage ESC members to continue to participate in an advisory capacity to the NIEM Open, PGB. Details of Open Project Sponsorship can be found online at "https://www.oasis-open.org/join-2/" and by contacting Dee Schur, the OASIS Open Source and Standards Advocate at dee.schur@oasis-open.org.

J. Clark Cully, Ph.D.
Deputy Chief Data Officer
Department of Defense

STANDARDS THAT INCLUDE NIEM

- ANSI APCO Alarm Monitoring Company to Public Safety Answering Point (PSAP)
- OASIS LegalXML Electronic Court Filing Technical Committee
- OASIS EXDL, Emergency Data Exchange Language FEMA (Emergency Management Domain)
- Nation Fire Protection NFPA 950 calls for compliance with NIEM. Emergency Incident Data Document (EIDD) – NISTIR 8255
- Model Minimum Uniform Crash Criteria Guideline (MMUCC) – Department of Transportation
- ANSI 42.42 Radiological Nuclear detectors – CBRN Domain
- NIST Big Data Framework Vol 7
- Biometrics ANSI/NIST ITL Standard – Biometrics Domain
- Biometric Conformance Test Software (BioCTS)
- BioCTS for AN-ITL v2 is a desktop application which tests electronic biometric data files, known as transactions, for conformance to NIST Special Publication (SP) 500-290
- Conformance Test Architecture (CTA) and Test Suite (CTS) called "BioCTS for AN-2011 NIEM XML" designed to test implementations of AN-2011 NIEM XML encoded transactions.
- NIEM cited in patent – Integrated Environment for Developing Information Exchanges patent No: US 8,769,480 B1 Dated July 1, 2014

DATA MANAGEMENT BOOK OF KNOWLEDGE (DMBOK)

1.3.8 Data Exchange Standards

Data Exchange Standards are formal rules for the structure of data elements. ISO (International Standards Organization) has developed data exchange standards, as have many industries. A data exchange specification is a common model used by an organization or data exchange group that standardizes the format in which data will be shared. An exchange pattern defines a structure for data transformations needed by any system or organization exchanging data. Data needs to be mapped to the exchange specification.

Although developing and agreeing on a shared message format is a major undertaking, having an agreed upon exchange format or data layout between systems can significantly simplify data interoperability in an enterprise, lowering the cost of support and enabling better understanding of the data.

The National Information Exchange Model (NIEM) was developed to exchange documents and transactions across government organizations in the United States. The intention is that the sender and receiver of information share a common, unambiguous understanding of the meaning of that information. Conformance to NIEM ensures that a basic set of information is well understood and carries the same consistent meaning across various communities, thus allowing interoperability.

NIEM uses Extensible Markup Language (XML) for schema definitions and element representation, which allows the structure and meaning of data to be defined through simple, but carefully defined XML syntax rules.

INTERNATIONAL USERS

Region	Countries
Asia	Japan
AUSTRALIA	Australia Defence Force
	New Zealand
EUROPE	European Commission's Interoperability Solutions for European Public Administrations (ISA) Programme ; Europol; Eurojust + 10 members; Belgium, Sweden
NORTH AMERICA	Canada (Employment and Social Development Canada (ESDC); Canada Border Services Agency; Public Safety Canada (PS); Citizenship and Immigration Canada (CIC), in cooperation with the Royal Canadian Mounted Police (RCMP) and Canada Border Services Agency (CBSA)); US/CANADA/MEXICO (CIOs of Canada, Mexico, and the United States)
SOUTH AMERICA	Columbia
MULTI - NATIONAL	United Nations (UNCEFACT), EU (JOINUP Project)

INTERAGENCY EXCHANGES BUILT ON NIEM

- **Interstate Justice and Public Safety Network (NLETS)** - exchange mission-critical law enforcement information.
- **Amber Alert** - child abduction emergency alert.
- **Prescription Drug Monitoring Exchange (PMIX)** - Mitigate pharmaceutical drug abuse across state lines.
- **FBI Incident Reporting National Data Exchange System (N-Dex)** standardized and secure criminal justice information sharing to relevant criminal justice agencies.
- **FBI National Crime Information Center (NCIC)** - United States' central database for tracking crime-related information.
- **Centers for Disease Control and Prevention (CDC) Emergency Preparedness and Response Exchange Requirements** - exchange and sharing of critical emergency management (EM) data.
- **National Electronic Interstate Compact Enterprise (NEICE)** - placement of children in foster care across state lines.
- **Indian Child Welfare Act (ICWA) e-Notice.**
- **LegalXML Electronic Court Filing** - XML-based standards support the implementation of electronic court filing

STANDARDS THAT INCLUDE NIEM

- ANSI APCO Alarm Monitoring Company to Public Safety Answering Point (PSAP)
- OASIS LegalXML Electronic Court Filing Technical Committee
- OASIS EXDL, Emergency Data Exchange Language FEMA (Emergency Management Domain)
- Nation Fire Protection NFPA 950 calls for compliance with NIEM. Emergency Incident Data Document (EIDD) – NISTIR 8255
- Model Minimum Uniform Crash Criteria Guideline (MMUCC) – Department of Transportation
- ANSI 42.42 Radiological Nuclear detectors – CBRN Domain
- NIST Big Data Framework Vol 7
- Biometrics ANSI/NIST ITL Standard – Biometrics Domain
- Biometric Conformance Test Software (BioCTS)
- BioCTS for AN-ITL v2 is a desktop application which tests electronic biometric data files, known as transactions, for conformance to NIST Special Publication (SP) 500-290
- Conformance Test Architecture (CTA) and Test Suite (CTS) called "BioCTS for AN-2011 NIEM XML" designed to test implementations of AN-2011 NIEM XML encoded transactions.
- NIEM cited in patent – Integrated Environment for Developing Information Exchanges patent No: US 8,769,480 B1 Dated July 1, 2014

DATA MANAGEMENT BOOK OF KNOWLEDGE (DMBOK)

1.3.8 Data Exchange Standards

Data Exchange Standards are formal rules for the structure of data elements. ISO (International Standards Organization) has developed data exchange standards, as have many industries. A data exchange specification is a common model used by an organization or data exchange group that standardizes the format in which data will be shared. An exchange pattern defines a structure for data transformations needed by any system or organization exchanging data. Data needs to be mapped to the exchange specification.

Although developing and agreeing on a shared message format is a major undertaking, having an agreed upon exchange format or data layout between systems can significantly simplify data interoperability in an enterprise, lowering the cost of support and enabling better understanding of the data.

The National Information Exchange Model (NIEM) was developed to exchange documents and transactions across government organizations in the United States. The intention is that the sender and receiver of information share a common, unambiguous understanding of the meaning of that information. Conformance to NIEM ensures that a basic set of information is well understood and carries the same consistent meaning across various communities, thus allowing interoperability.

NIEM uses Extensible Markup Language (XML) for schema definitions and element representation, which allows the structure and meaning of data to be defined through simple, but carefully defined XML syntax rules.

INTERNATIONAL USERS

Region	Countries
Asia	Japan
AUSTRALIA	Australia Defence Force
	New Zealand
EUROPE	European Commission's Interoperability Solutions for European Public Administrations (ISA) Programme ; Europol; Eurojust + 10 members; Belgium, Sweden
NORTH AMERICA	Canada (Employment and Social Development Canada (ESDC); Canada Border Services Agency; Public Safety Canada (PS); Citizenship and Immigration Canada (CIC), in cooperation with the Royal Canadian Mounted Police (RCMP) and Canada Border Services Agency (CBSA)); US/CANADA/MEXICO (CIOs of Canada, Mexico, and the United States)
SOUTH AMERICA	Columbia
MULTI - NATIONAL	United Nations (UNCEFACT), EU (JOINUP Project)

TOOLS

- **CMF**: The NIEMOpen **Common Model Format** is a technology neutral data modeling specification being developed by the NTAC. It is currently supported by the CMF Tool, which is a command-line tool for the development of NIEM-conforming data exchange specifications using the NIEM Common Model Format (CMF).
  - **Benefits**
    - Supported by free and open-source tools which convert CMF data models into technology-specific developer artifacts: XML Schema documents, JSON Schema documents, etc.
    - Technology agnostic
  - **Status:** In progress- to be released with version 6.0 tools. Functionality being tested by developers and reporting feedback at NTAC meetings. CMF will be incorporated into NIEM API 2.0.

- **MEP**: The NIEMOpen Message Exchange Package (MEP) Builder is an open-source tool under development to guide users through the stages of the lifecycle and provide the necessary artifacts to develop a message exchange specification or IEPD.
  - **Benefits**
    - Simplifies the creation of message exchange packages
    - Intuitive UI/UX wizard guides the user step by step
      - Allows users to publish, browse, and modify IEPD's/MEP's
      - Open-source development allows for NIEM community to enhance/expand functionality.

TOOLS (CONT.)

- **API 2.0:** Application Programming Interface is in progress and will be used for:
  - Search and Discovery
  - Working on NIEM models
  - Transforming Models (CMF dependency)
  - Validation
  - Providing schema to external tool such as MEP.

– **Status:** In progress and to be completed as part of the NIEMOpen v6.0 toolset

- **Repos:** The transition of repositories from the NIEM GitHub account to the new NIEMOpenGitHub account is in -progress

NIEM Next Steps

IN-DEPTH DATA STEWARDS GROUP PRESENTATION

MONTHLY NIEM ONLINE TRAINING

AGENCY ADOPTION

Multi-State Data Literacy Research

Jamie Kimes

odga

5/31/2024

Colorado Data Literacy Research

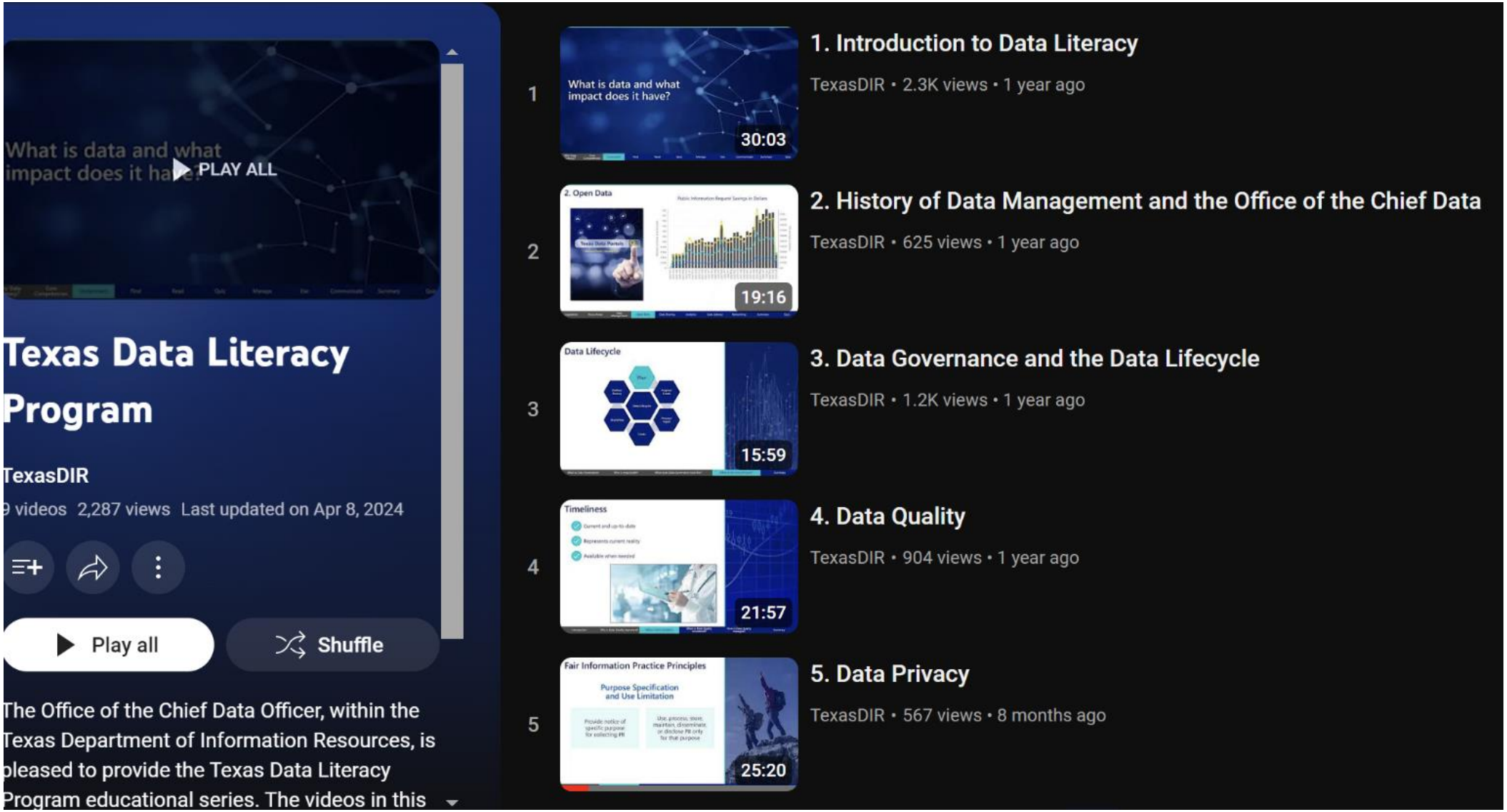
Questions:

	Percentage of Interviewed States Answering "Yes"
Is the state IT Centralized?	100%
Name of Centralized IT (link)	
Is there a Chief Data Officer?	92%
Is there a Chief Data Office?	62%
Name of Chief Data Office? (link)	
Is the Chief Data Office(r) Under the Centralized IT?	38%
Does the Chief Data Office Report to the Governor?	31%
Is there a Data Literacy Program?	46%
Is data literacy a documented strategic priority?	54%
Is the Data Literacy Program searchable from the centralized IT website?	15%

Questions:

	Percentage of Interviewed States Answering "Yes"
What Modalities are used:	
Public Virtual	8%
Public Documents	8%
Public PowerPoints	8%
LMS Videos	8%
Public Videos	23%
Access-Restricted Materials	15%
Is the training designed for specific roles?	15%
Is certification offered?	23%

Texas


What is data and what impact does it have?
PLAY ALL
Texas Data Literacy Program
TexasDIR
9 videos 2,287 views Last updated on Apr 8, 2024
Play all
Shuffle
The Office of the Chief Data Officer, within the Texas Department of Information Resources, is pleased to provide the Texas Data Literacy Program educational series. The videos in this
1
What is data and what impact does it have?
30:03
1. Introduction to Data Literacy
TexasDIR • 2.3K views • 1 year ago
2
2. Open Data
19:16
2. History of Data Management and the Office of the Chief Data
TexasDIR • 625 views • 1 year ago
3
Data Lifecycle
15:59
3. Data Governance and the Data Lifecycle
TexasDIR • 1.2K views • 1 year ago
4
Timeliness
21:57
4. Data Quality
TexasDIR • 904 views • 1 year ago
5
Fair Information Practice Principles
Purpose Specification and Use Limitation
25:20
5. Data Privacy
TexasDIR • 567 views • 8 months ago

Indiana

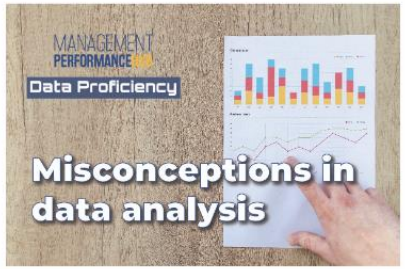
Level 1: Green Badge Series

Lesson 1 - What is Data and Data Literacy?

Whether it is filling out a customer intake form, using restaurant reviews to decide your dinner destination or reviewing the latest dashboard at coronavirus.IN.gov, each of us uses data on a daily basis to learn and guide our decision-making. Often, we interact with data without realizing it.

Go to Lesson 1

Lesson 2 - Misconceptions in Data Analysis



Statistics are all around us every day. Most sources present these statistics as fact, but what do they really mean? Do they give you all the information you need to know?

Go to Lesson 2

Maryland

Training

The State is currently developing three training courses that will be available to state employees:

- Introduction to Data Management
- Effective Data Stewardship
- Data for the Executive Leader

*These courses will be made available shortly to State of Maryland employees.

Maryland

Resources

What is Data Literacy?

Data is being constantly created and consumed. Data literacy is the ability to interpret, analyze, and think about this data. To make informed decisions about a problem an individual relies on their data literacy to dig into this information. Thinking critically while reading and analyzing data is vital in this process. By doing this, the data consumer for themselves discerns the story the data presents. Being data literate empowers consumers to make data informed decisions rather than making decisions based on emotion or the whims of others.

Watch Video

Oregon

OREGON DATA LITERACY FRAMEWORK REPORT

October 2023

Oregon

FOUNDATIONAL DATA LITERACY

An individual at the foundational level of data literacy will demonstrate the following skills/abilities in three core elements:

Understanding

- Basic awareness of concepts and techniques

Ability

- Follows guidance

Interaction with Others

- Seeks advice

It is recommended that the following checked roles achieve a foundational level of data literacy in the following capabilities to be most effective

	Decision Makers	Organizational Data Leads	Data Managers	Subject Matter Experts	Data Architects/Engineers/ Administrators	Data Analysts	General Data Users	Data Collectors/Data Entry Staff	Data Communicators
Value Data Assets	☑	☑	☑	☑	☑	☑	☑	☑	☑
Collection	☐	☑	☑	☑	☑	☑	☑	☑	☑
Documentation	☐	☑	☑	☑	☑	☑	☑	☑	☑
Organization/Management	☐	☑	☑	☑	☑	☑	☐	☑	☑
Privacy & Security	☑	☑	☑	☑	☑	☑	☑	☑	☑
Governance	☑	☑	☑	☑	☑	☑	☐	☐	☑
Modeling	☐	☑	☐	☐	☑	☑	☐	☐	☐
Integration	☐	☑	☑	☑	☑	☑	☑	☐	☑
Cleaning	☐	☑	☑	☑	☑	☑	☐	☑	☐
Quality Evaluation	☑	☑	☑	☑	☑	☑	☑	☑	☑
Discovery	☐	☑	☑	☑	☑	☑	☑	☐	☑
Analysis	☐	☑	☑	☑	☑	☑	☑	☐	☑
Visualization	☐	☑	☑	☑	☑	☑	☑	☐	☑
Interpretation	☑	☑	☑	☑	☑	☑	☑	☐	☑
Storytelling	☑	☑	☑	☑	☑	☑	☑	☐	☑
Evidence-Based Decision-Making	☑	☑	☑	☑	☑	☑	☑	☑	☑
Ethics	☑	☑	☑	☑	☑	☑	☑	☑	☑

Oregon

INTERMEDIATE DATA LITERACY

An individual at the intermediate level of data literacy will demonstrate the following skills/abilities in three core elements:

Understanding

- Demonstrates a broad understanding of concepts and techniques

Ability

- Demonstrates application of concepts and techniques with minimal guidance in normal situations
- Uses precedents and/or industry standards

Interaction with Others

- Influences, upholds, modifies, consults others with more experience
- Shares advice

It is recommended that the following checked roles achieve an intermediate level of data literacy in the following capabilities to be most effective:

	Decision Makers	Organizational Data Leads	Data Managers	Subject Matter Experts	Data Architects/Engineers/ Administrators	Data Analysts	General Data Users	Data Collectors/Data Entry Staff	Data Communicators
Value Data Assets	☑	☑	☑	☐	☑	☑	☐	☐	☑
Collection	☐	☑	☑	☑	☐	☑	☐	☐	☐
Documentation	☐	☐	☑	☑	☑	☑	☐	☐	☐
Organization/Management	☐	☐	☑	☑	☑	☑	☐	☐	☐
Privacy & Security	☐	☐	☑	☑	☑	☑	☐	☐	☐
Governance	☐	☐	☑	☑	☑	☐	☐	☐	☐
Modeling	☐	☐	☐	☐	☑	☑	☐	☐	☐
Integration	☐	☑	☑	☐	☑	☑	☐	☐	☐
Cleaning	☐	☑	☐	☑	☑	☑	☐	☐	☐
Quality Evaluation	☐	☑	☑	☑	☑	☑	☐	☐	☑
Discovery	☐	☐	☑	☑	☑	☑	☐	☐	☑
Analysis	☐	☐	☐	☐	☐	☑	☐	☐	☐
Visualization	☐	☐	☐	☐	☐	☑	☐	☐	☑
Interpretation	☑	☑	☑	☐	☑	☑	☐	☐	☑
Storytelling	☐	☑	☐	☑	☑	☑	☐	☐	☑
Evidence-Based Decision-Making	☑	☑	☑	☑	☐	☑	☐	☐	☑
Ethics	☐	☐	☑	☑	☑	☑	☐	☐	☑

Oregon

ADVANCED DATA LITERACY

An individual at the advanced level of data literacy will demonstrate the following skills/abilities in three core elements:

Understanding

- Demonstrates an extensive understanding of concepts and techniques

Ability

- Demonstrates extensive application of concepts and techniques
- Uses industry standards and sometimes modifies uses of precedents
- Shapes the organization's approach in the application of this skill/knowledge area

Interaction with Others

- Leads, implements, monitors, regulates, advises, exp

It is recommended that the following checked roles achieve an advanced level of data literacy in the following capabilities to be most effective

	Decision Makers	Organizational Data Leads	Data Managers	Subject Matter Experts	Data Architects/Engineers/ Administrators	Data Analysts	General Data Users	Data Collectors/Data Entry Staff	Data Communicators
Value Data Assets	☐	☑	☐	☐	☑	☐	☐	☐	☐
Collection	☐	☐	☐	☐	☐	☐	☐	☐	☐
Documentation	☐	☑	☐	☐	☑	☐	☐	☐	☐
Organization/Management	☐	☑	☐	☐	☑	☐	☐	☐	☐
Privacy & Security	☐	☑	☐	☐	☑	☐	☐	☐	☐
Governance	☐	☑	☐	☐	☐	☐	☐	☐	☐
Modeling	☐	☐	☐	☐	☑	☐	☐	☐	☐
Integration	☐	☐	☐	☐	☑	☐	☐	☐	☐
Cleaning	☐	☐	☐	☐	☑	☐	☐	☐	☐
Quality Evaluation	☐	☐	☐	☐	☑	☑	☐	☐	☐
Discovery	☐	☑	☐	☐	☑	☑	☐	☐	☐
Analysis	☐	☐	☐	☐	☐	☑	☐	☐	☐
Visualization	☐	☐	☐	☐	☐	☑	☐	☐	☐
Interpretation	☐	☐	☐	☐	☐	☑	☐	☐	☐
Storytelling	☐	☐	☐	☐	☐	☑	☐	☐	☑
Evidence-Based Decision-Making	☐	☐	☐	☐	☐	☐	☐	☐	☐
Ethics	☐	☑	☐	☐	☐	☐	☐	☐	☐

Problems Data Literacy Might Solve

The business side often asks for things like a "data warehouse" or a system that connects to another using an API, but they may not fully understand what they're requesting.
Business is "ripe for vendor picking" because they don't understand how to start with the problem that needs to be solved and to use existing tools.
If a division-director level person is presenting to the JBC and gets asked a data question, they can't address it in real time.
Data people having to spend a ton of time repeating themselves.
Executive Leadership wants reports but doesn't understand they need the infrastructure to enable those reports. Executive Leadership regurgitates what the data team tells them but doesn't truly understand it.
We present data, but no one acts on it because they don't understand how to turn it into action.

Colorado Progress

- Have begun translating ASU-type material into public sector language
- Have begun developing a course syllabus for foundational data literacy
- Have leveraged data from data inventory to identify possible communities of practice

COV Training Prioritization

1. Data Privacy
2. Data Quality
3. Data lifecycle
4. Understanding Data
5. Introduction to Data Management
6. Effective Data Stewardship
7. Data for the Executive Leader
8. Data Literacy
9. Other

- Best format?
  - videos, quizzes, lunch and learns, etc.
- Length?
- Required Training (Y/N)?

ODGA Recommendations
1. Data Quality
2. Data for the Executive Leader
3. Data Literacy

Motion Needed

Unstructured Data Scanning

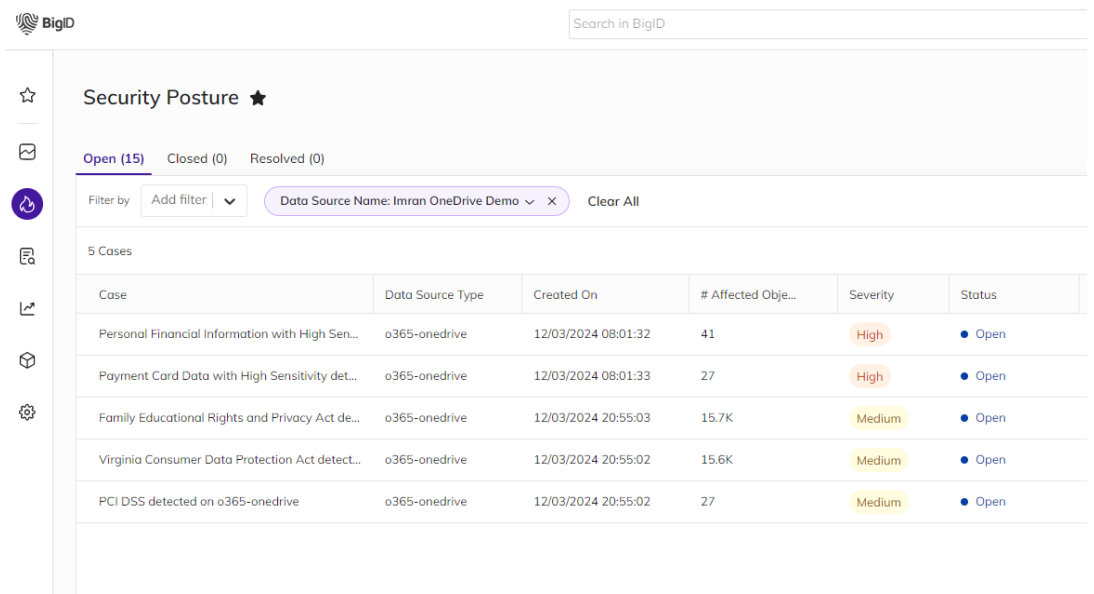
Chris Burroughs and Imran Afridi

odga

Reducing Risk through Unstructured Data Scanning

Identify Files with:

- Credit Cards
- Passwords in Clear Text
- Drivers License Numbers
- Tax ID's
- Passport Numbers
- PII
- And custom categories



Pilot Results Show the Unidentified Risk

Clear text passwords & keys

Credit card numbers

Driver's license

Social Security numbers

Bank account numbers

Passport numbers

ODGA Update

Marcus Thornton and Chris Burroughs

Datathon 2024

City of Norfolk

Alpha1Data (General Public)

Data Tribe (W&M)

W&M MSBA

People's Choice

Best Datathon ever!

The Communications team is currently working to see how the Commonwealth can implement some Datathon solutions!








Solutions

- 1st Place: A data informed employment notification program.
- 2nd Place: "Commonwealth AI Career Hub Portal"
- 3rd Place: WorkWatch VA, an interactive Power BI dashboard with three main components: Virginia Employment Statistics, Unemployment rate forecast, and a Policy Impact estimator
- People's Choice: A model was created based on labor force participation and demographics such as race.

Datathon 2024

- **Participating Organizations:**

Public Sector	Private Sector	Universities	High School
VDH City of Norfolk DMAS DCJS	Capital One Dominion Energy Pave Finance VIO LLC AWS Federal Reserve Bank of Richmond	William and Mary Virginia Military Institute George Mason University Virginia Commonwealth University	JR Tucker High School

- **Participating Sponsors:** OpenGov, Voyatek (GCOM), Vertica, CarMax, Kearney and Company, Dataversity, and Macon IT, Inc.
- **Data Providers:** VEC, SCHEV, SBSD, DHRM, VCCS, VA Works, DARS, DOE

Data Management Maturity Assessment

Results and Objectives

Data Management Maturity Assessment

Measures agency progress in: People, Processes, Businesses, Technology

Allows ODGA to determine how to best assist agencies through services and trainings

Allows ODGA to connect agencies to ODGA services that ultimately help Virginians live better

Response Rate Breakdown

Sent to **104** Organizations, Responses from **63**

52/73
Agencies

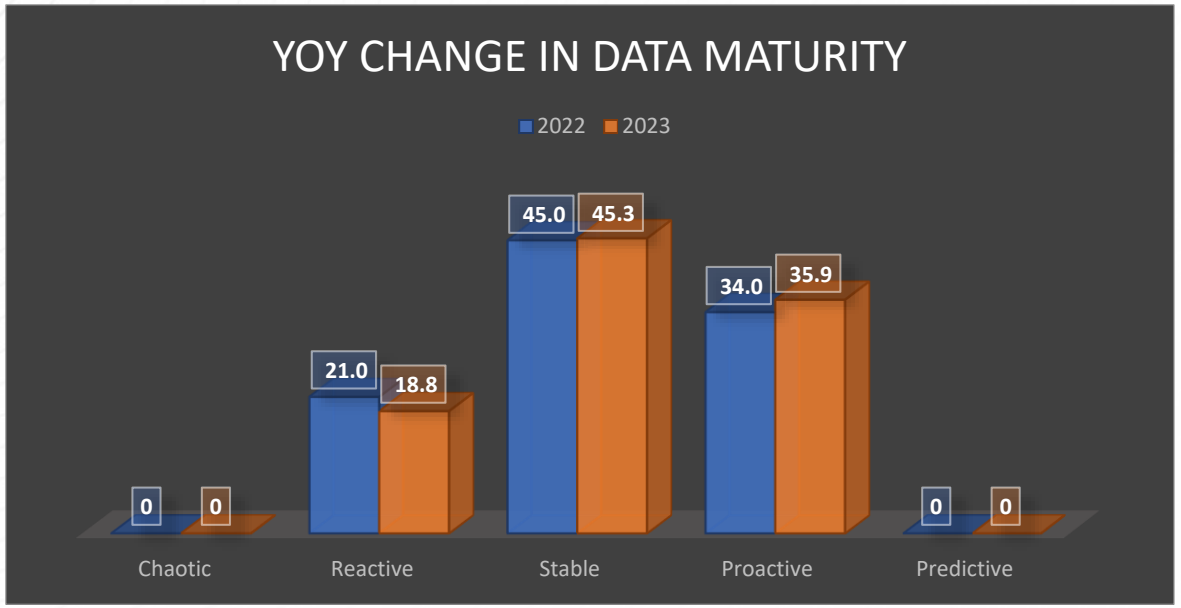
2/5
Gov's Office Orgs.

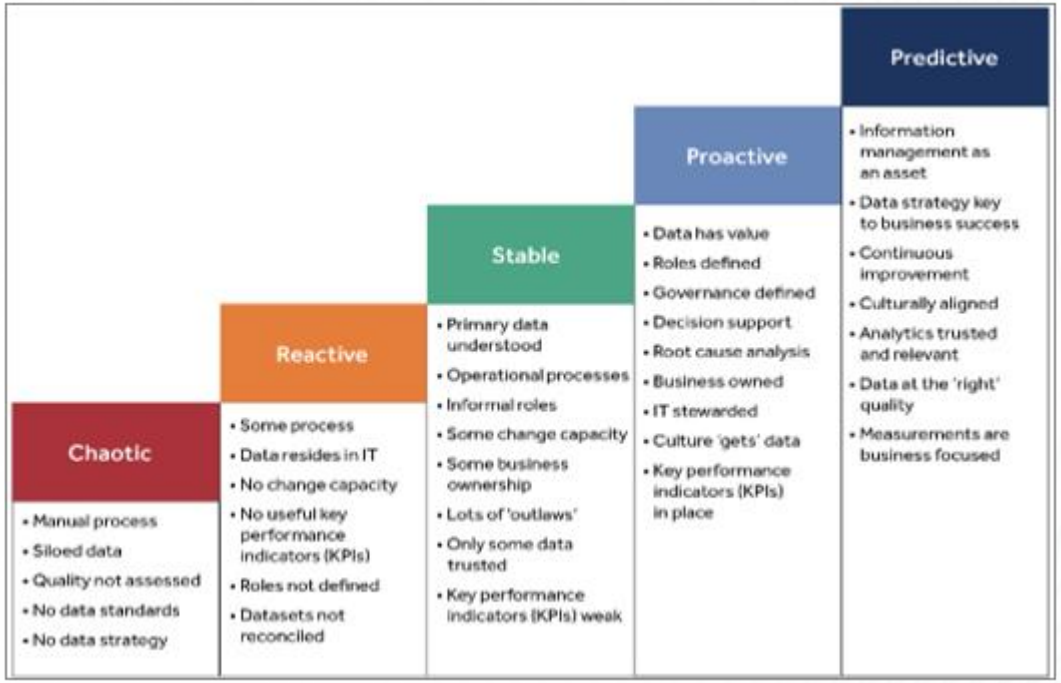
3/5
Museums

5/9
Institutes of Higher Ed

1/12
Other

Slight Improvement in Data Maturity





- Incremental progress made as less agencies rated reactive and more moving to a proactive status
- Overall COV average score increased from **2.48** to **2.56** YOY

Building On Strengths

Roles for Existing Data Activities

- 88% of agencies (56) have formal roles for data established, with data stewards and owners.
- ODGA established a resource library with role definitions, RACI, and job descriptions to help agencies mature in this area

Management Support

- Senior Management Sponsorship for data initiatives at agencies. Average Score = 3.09. Only one agency rated themselves a "1."
- ODGA established the Executive Data Board and Data Governance Council to strengthen management support

ODGA Can Help Address Assessment Gaps

Key Focus Areas	Solutions
Approach/Plan for Improving Data Quality 28.13% rated maturity level 1	•Leverage ODGA's Data Quality tool to baseline and improve quality •Use ODGA Data Quality Standards and Data Entry Standards templates •Expand Data Quality Awareness education within agencies
Lack of Metrics on Data Quality	
Metadata Management Overall score – 1.89 of 4.0; 8 agencies rated as 0	•Deploy NIEM Standard and EA 225 guidance •Use ODGA's Purview tool for metadata management •Contract with ODGA for consulting assistance
Data Modeling 59% of agencies (38) have trouble with data modeling. 16% regard their data as being primarily in silos.	•Contract with ODGA for consulting assistance on data modeling and data lineage
Use of Logs to Record Data Management Risks/Issues 31% of agencies rated "1" or lower	•Use the risk log template available from ODGA •Enter data risks into VITA's Archer system

Next Steps

Deliver Recommendations. Provide agency-specific governance “prescriptions” with action items for improvement in key areas.

Determine desired outcomes. Evaluate business objectives, existing outcomes and processes for determination as a larger part of top 5 key outcomes..

Determine Datasets. Develop COV data catalog.

Facilitate Sharing. Identify datasets and sources needed to facilitate objectives. Incorporate into Commonwealth DataTrust.

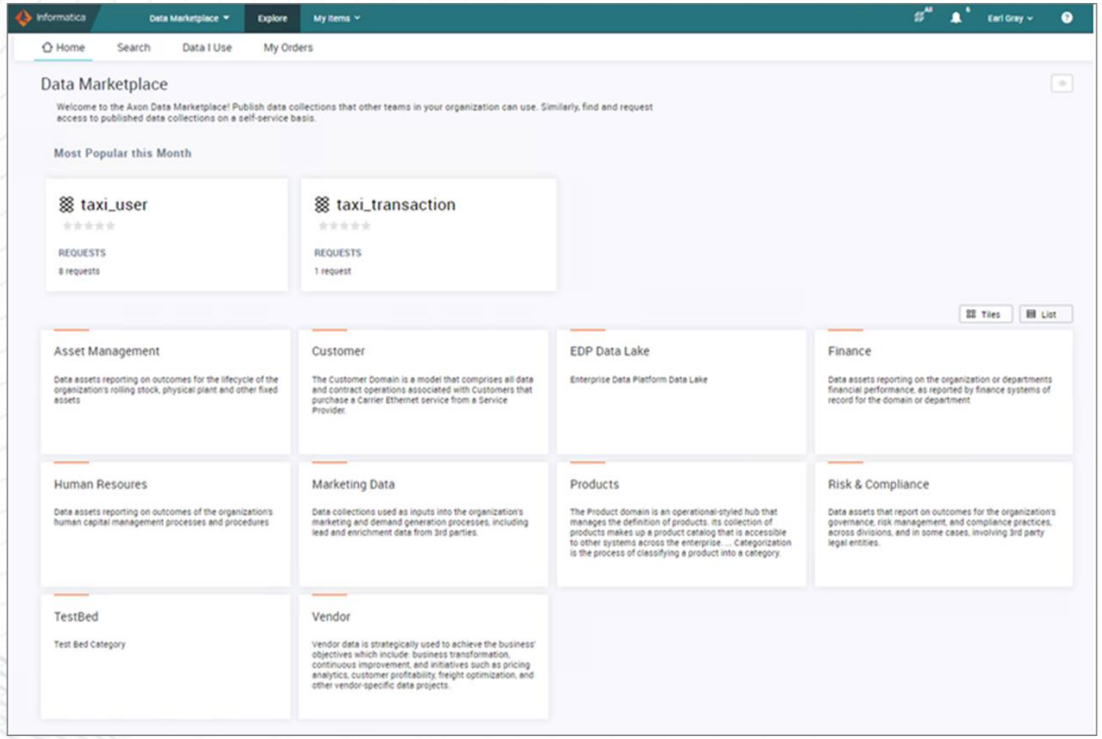
ODGA Data Catalog

ODGA is building a Commonwealth-wide catalog of data assets to allow agencies to easily find and access trusted data so they may garner data-driven insights and perform in-depth analysis that was never possible before.

Benefits:

- Better visibility into your agency's data
- Data democratization
- New insights
- Faster decision-making
- Better data governance
- Reduced risk through improved regulatory compliance

ODGA Data Catalog



Examples:

- A list of your customers or clients and the attributes related to them
- Financial Information like budgets and expenses
- A list of state parks or other state assets
- A database of licenses or permits
- A list of 3rd party providers and their corresponding information

- A library or **inventory** of all your data sets, visualizations, and dashboards.
- Uses **metadata** combined with **data management** and **search tools** to help agencies manage their data

IMPORTANT

This catalog will only be available for Commonwealth of Virginia agencies. **It will not be published to citizens**

If you'd like to share your data with Virginia citizens, leverage the Open Data Portal data.virginia.gov

ODGA Resource Library

FREE resources on the ODGA Website for agencies. Includes templates, videos, guidebooks, etc!

- ODGA at a Glance
- Data Governance RACI Template
- Risk Assessment Template
- Data Risk Register Template
- Data Governance Roles
- Data Governance Roles- Posters
- Data Governance Job Description Samples
- Data Strategy Example
- Data Governance Guidebook
- Data Governance Council Guidebook
- Protecting Unstructured Data Guidebook
- Protecting Structured Data Guidebook
- BI Reporting Governance Guidebook
- Data Camp Data Storytelling Cheat Sheet
- Data Camp Data Governance Cheat Sheet
- Data Camp Data Quality Dimensions Guide
- ASU Data Literacy YouTube Course

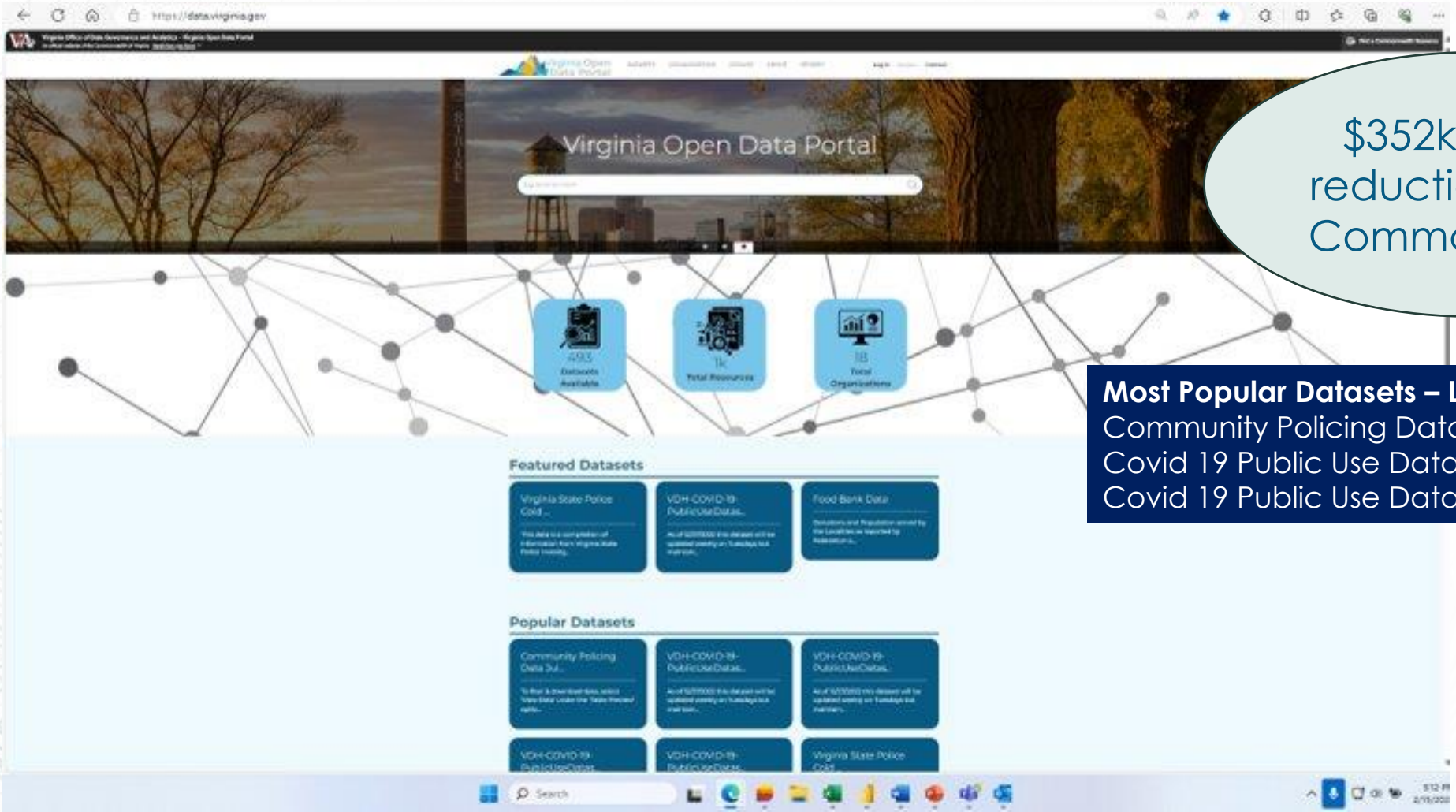
Most Popular

ODGA Project Updates

ODGA Projects

- DOF Forestry Project
  - Improves accuracy, Reduces labor requirement, Eliminates a single point of failure, Increases transparency into $'s by county
- Virginia Community Engagement Index
  - Saved up to **$45,000 versus an external vendor,** Provided expertise in Power BI that the Serve Virginia Team did not have, Saved money they would have needed to spend for ad hoc questions with another vendor.
- DFS Case Statistic Dashboards
  - Improved accuracy due to elimination of manual processes, Reverse engineered complex stored procedures and provided clarity into how they operate to DFS internal staff, Saves DFS team 3 hours monthly/ versus manual effort.
- Virginia Permit Transparency
  - Developed 5 dashboards that pulled aggregated data from a VITA data warehouse, Deliverables provided **one month early, Cost avoidance up to $200,000 for the VPT team versus an external vendor.**

New and Improved Open Data Portal



$352k/yr cost reduction to the Commonwealth

Most Popular Datasets – Last 30 Days
Community Policing Data – 592 views
Covid 19 Public Use Dataset – Cases – 455 views
Covid 19 Public Use Dataset – Zipcode – 139 views

Executive Agency Data Trust Status

Administration	Agriculture	Commerce	Education		Finance	Health & HR	Labor	Natural Resources	Public Safety	Transportation	Veterans
CB	**DOF**	DHCD	CNU	**SCHEV**	BOA	ATLFA	**DOLI**	**DCR**	CASC	DMV	DMA
DGS	VDACS	**DSBSD**	**DOE**	SMV	DOA	CCCA	DPOR	DEQ	**DCJS**	**DOAV**	**DVS**
DHRM	VRC	ENERGY	EVMS	SVHEC	DPB	CH	**VEC**	DHR	**DFP**	DRPT	VSF
ELECT		TRRC	FCMV	SWHEC	TAX	CSH		**DWR**	**DFS**	MVDB	
VITA		**VEDP**	GH	UMW	**TRS**	CVTC		MRC	**DJJ**	OIPI	
ODGA		**VIPC**	GMU	UVA	VRA	**DARS**			**DOC**	VAP3	
		VTA	IALR	VCA		**DBHDS**			**VDEM**	VCSFA	
			JMU	**VCCS**		**DBVI**			VPB	VDOT	
			JYF	**VCU**		DHP			**VSP**	VPA	
			LVA	VIMS		**DMAS**				VPRA	
			LWU	VMFA		**DSS**					
			NCI	VMI		ESH					
			NSU	VMNH		HDMC					
			ODU	VSDB		NVMHI					
			RBC	VSU		**OCS**					
			RHEA	VT		PGH					
			RU	WM		SEVTC					
						SVMHI					
						SWVMHI					
						VBPD					
						VCBR					
						VDDHH					
						VDH					
						VFHY					
						VRCBVI					
						WSH					
						WWRC					

Other Members

OSIG
Virginia 529
DCLS
HSD
JCHC
OMNI Institute
VARR
VHHA
CTO
G H International
Qlarion
Rappahannock CSB

Chesapeake PD
Chesterfield PD
Danville PD
Emporia PD
Hampton PD
Hopewell PD
Lynchburg PD
Martinsville PD
Newport News PD
Norfolk PD
Petersburg PD
Portsmouth PD
Richmond PD
Roanoke PD
Virginia Beach PD

Agencies in Bold have signed either Member or User Commonwealth Data Trust Agreements

Data Quality Profiling

Rebecca Armstrong

odga

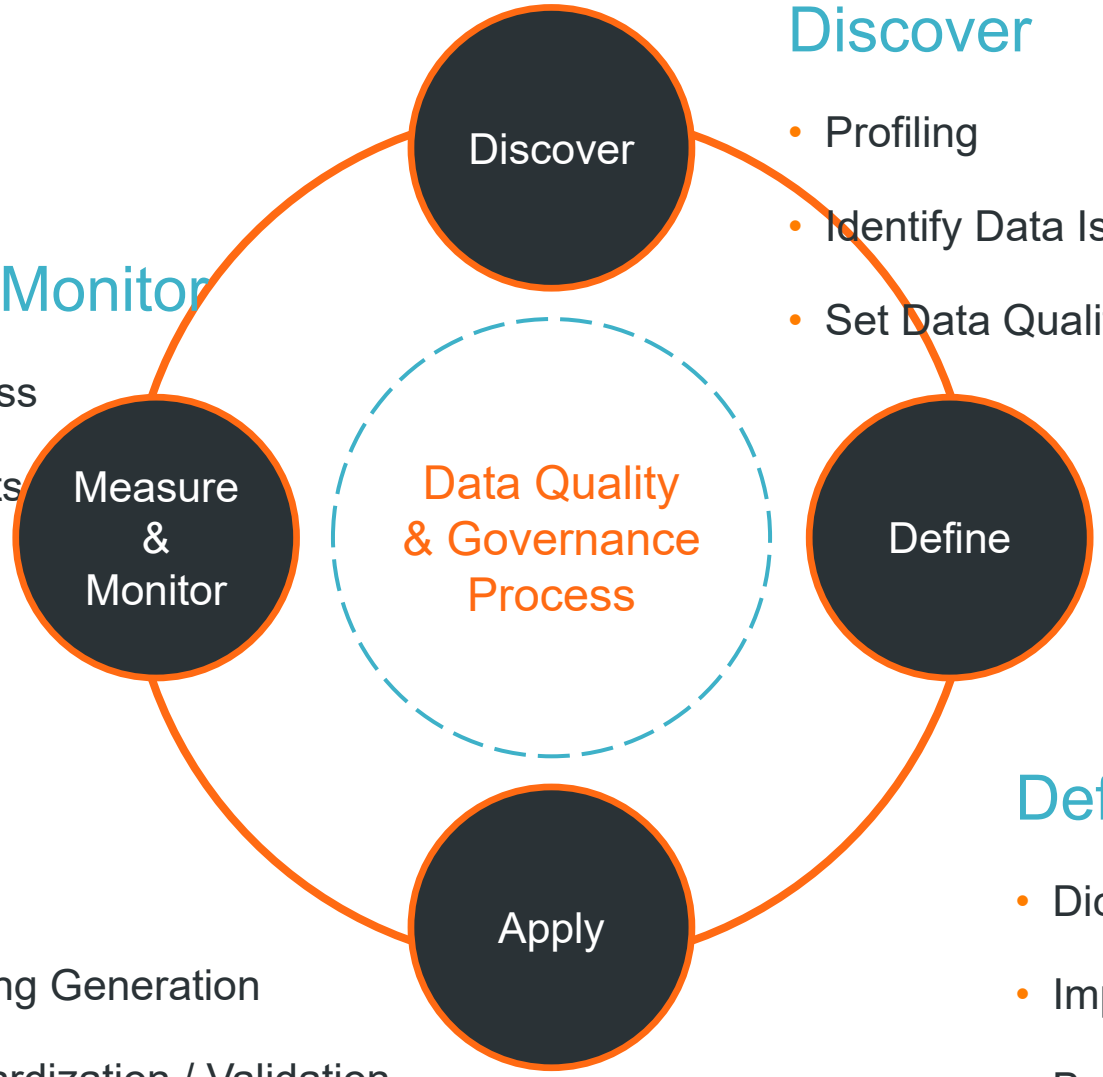
5/31/2024

Informatica

Informatica Data Quality for VA Office of Data Governance and Analytics

Where data & AI come to LIFE

Data Quality Methodology



Discover

- Profiling
- Identify Data Issues
- Set Data Quality Goals

Define

- Dictionaries
- Implement Rules
- Build Cleanse, Parse,

Apply

- Mapping Generation
- Standardization / Validation
- Matching / Consolidation

Measure & Monitor

- Review Progress
- Threshold Alerts
- Scorecards

ODGA Enable Agencies to Assess Data Quality

Enable agencies to profile specific tables to get granular details about their data

Discover

- Profiling
- Identify Data Issues
- Set Data Quality Goals

Create iterative data quality monitoring with scorecards

Measure & Monitor

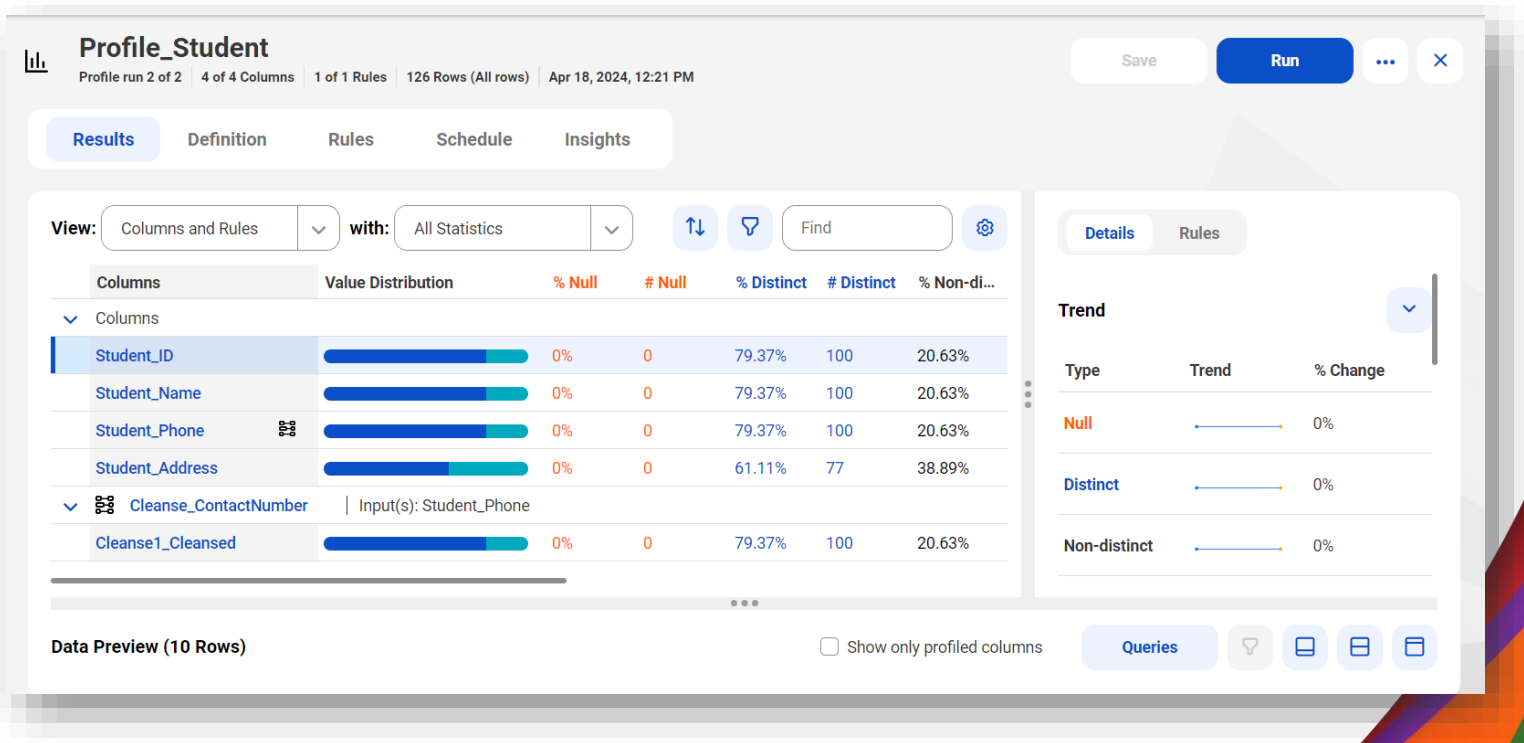
- Review Progress
- Threshold Alerts
- Scorecards

Discover Data Anomalies

Profile data to examine its structure and context using out-of-the-box templates

Drill down to see details and filter on results

Compare profile runs to identify trends over time

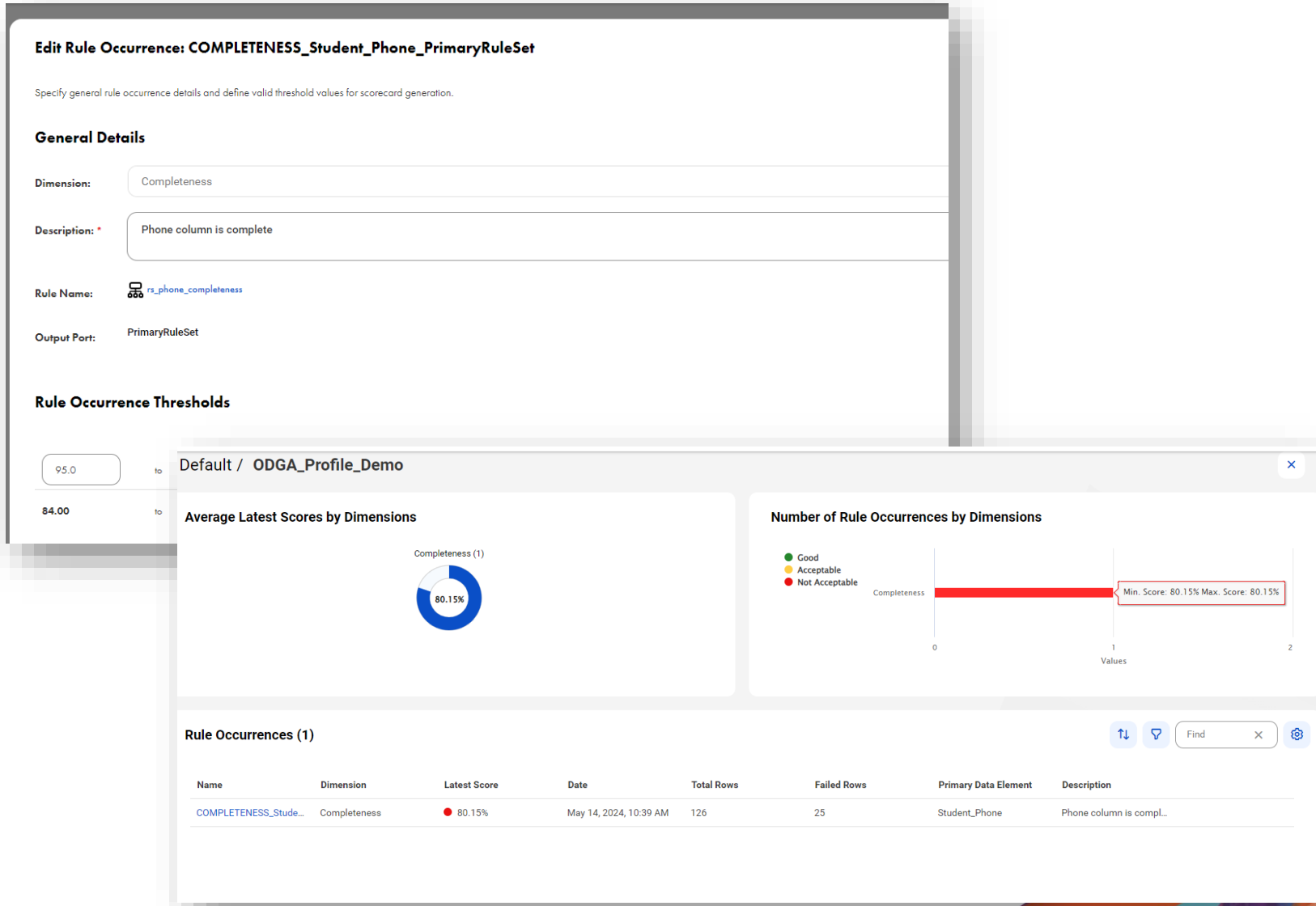


Measure & Monitor

Associate rule occurences with profiles

Build score cards to monitor how quality changes over time

Set thresholds and alerts for when quality dips



Questions?