Why it's time to revisit **Data Governance**

McCourt School
MASSIVE DATA INSTITUTE

Dr. Jacob M. Pasner
Assistant Research Professor
COPAFS Quarterly Meeting
December 5, 2025

Why does this matter right now?

45 47

PRESIDENTIAL ACTIONS

Stopping Waste, Fraud, and Abuse by Eliminating Information Silos

Executive Orders | March 20, 2025

Assessing the Health of the Federal Statistical Agencies

The Nation's Data at a Crossroads

Year Two Status Report

45 47

PRESIDENTIAL ACTIONS

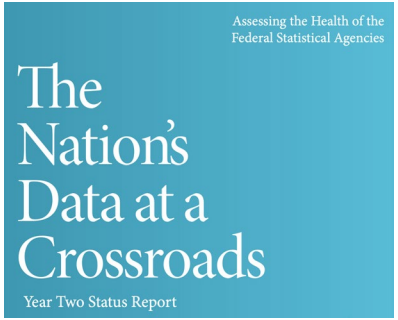
REMOVING BARRIERS TO AMERICAN LEADERSHIP IN ARTIFICIAL INTELLIGENCE

The White House | January 23, 2025

Why does this matter right now?

Unlocking Data Access

Translating Data into Public Good



Data Generation

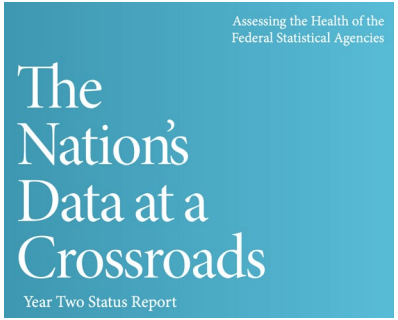
Data Governance → Data Access

Data-Driven Innovation

Why does this matter right now?

Unlocking Data Access

Translating Data into Public Good



Data Generation

Data Governance → Data Access

Data-Driven Innovation

What is…

Data Governance?

Data Governance

Is the set of principles and policies that empower an agency to ensure that:

1. Data is of good quality and fit for permitted use;

Data Governance

Is the set of principles and policies that empower an agency to ensure that:

1. Data is of good quality and fit for permitted use;
2. Data is secure and privacy is protected;

Data Governance

Is the set of principles and policies that empower an agency to ensure that:

1. Data is of good quality and fit for permitted use;
2. Data is secure and privacy is protected;
3. Data is Findable, Accessible, Interoperable, and Reproducible (FAIR);

Data Governance

Is the set of principles and policies that empower an agency to ensure that:

1. Data is of good quality and fit for permitted use;
2. Data is secure and privacy is protected;
3. Data is Findable, Accessible, Interoperable, and Reproducible (FAIR);
4. Data is in compliance with industry and regulatory standards;

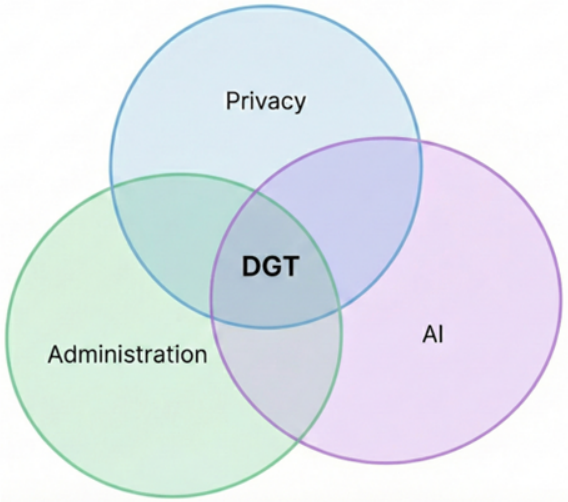
Data Governance

Is the set of principles and policies that empower an agency to ensure that:

1. Data is of good quality and fit for permitted use;
2. Data is secure and privacy is protected;
3. Data is Findable, Accessible, Interoperable, and Reproducible (FAIR);
4. Data is in compliance with industry and regulatory standards;
5. Cloud computing and contracting costs are kept to a minimum.

"Data Governance" covers the tangle of data, systems, policy and people we need to navigate to select, confirm and deploy the solution. Data Governance Transformation is how organizations make sure they're **prepared** to successfully execute data projects.

Data **Governance Transformation** is the building block for any data project



Privacy problem specification tasks must be done for **all** participating partners at the ***project outset***. All specification tasks are completed by members of the Data Governance chain of command of the participating partners.

Data Project Specification Checklist:

- **Data Categorization:** What is the sensitivity level of information as determined by agency policy and terms specified during collection.
- **RBAC:** Who is allowed to see/manipulate/grant access for any given data that will be implicated. This may include upstream data objects.
- **Data Architecture:** What data standards/validations/quality assurance/Primary Keys are implemented in the implicated databases.
- **IT Infrastructure:** Where is the data stored and what will need to be connected together to allow Authenticated users to do the required work.
- **Privacy Policy Requirements:** Which Privacy Policy regulations will be implicated for these data, their categorization, permitted uses, and the identified users.
- **Chain of Command:** Who is allowed to make decisions about allowable uses. Who will curate the data. Which stewards will be required to validate proper data usage.
- **Goal:** What is the expected outcome for all partners and how will this goal fulfill obligations/requirements/needs for them all (e.g. Permissible use for public good).
- **Transparency:** How can all participants (including people whose data is being used) keep up to date on project progress and outcomes to ensure proper usage and re-use of final data product.

Who cares about Data Governance standardization?

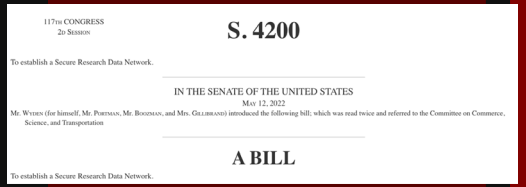
Policymakers

Foundations for Evidence -Based Policymaking Act

Outcomes-Driven Legislation

Serving Constituents

Responding to Emergencies



117TH CONGRESS
2D SESSION

S. 4200

To establish a Secure Research Data Network.

IN THE SENATE OF THE UNITED STATES
MAY 12, 2022

Mr. WYDEN (for himself, Mr. PORTMAN, Mr. BOOZMAN, and Mrs. GILLIBRAND) introduced the following bill; which was read twice and referred to the Committee on Commerce, Science, and Transportation

A BILL

To establish a Secure Research Data Network.

Agency Staff

Program Implementation

Efficiency and Cost Savings

Organizational Chain of Command

Clarifying Roles and Responsibilities

Academics

Research for Policy

Government Support

Understanding our World

McCourt School
MASSIVE DATA INSTITUTE

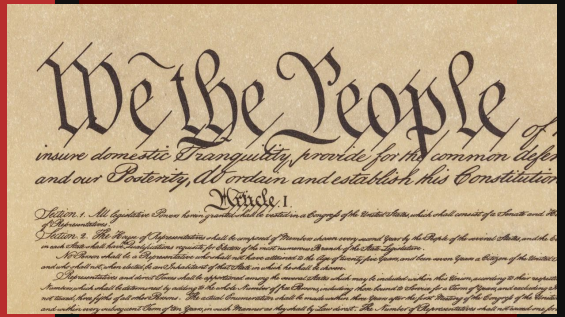
The Public

Build Social Trust

Deliver Timely Information and Services

Easy-to-use Interfaces

Leverage Open Data to Drive Innovation and Growth



We the People

insure domestic Tranquility, provide for the common defe
and our Posterity, do ordain and establish this Constitution

Article. I.

Section. 1. All legislative Powers herein granted shall be vested in a Congress of the United States, which shall consist of a Senate and
of Representatives.

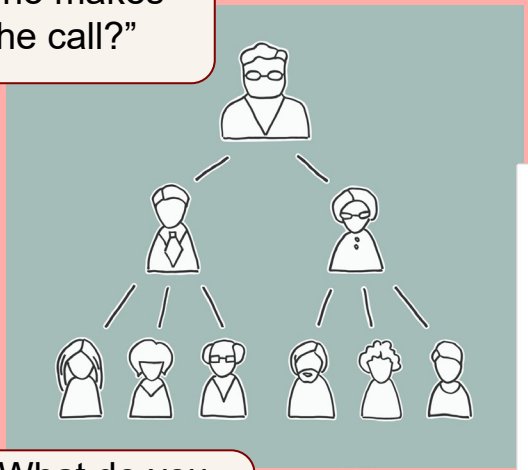
Section. 2. The House of Representatives shall be composed of Members chosen every second Year by the People of the several States, and the
and the Electors in each State shall have the Qualifications requisite for Electors of the most numerous Branch of the State Legislature.

No Person shall be a Representative who shall not have attained to the Age of twenty five Years, and been seven Years a Citizen of the United
and who shall not, when elected, be an Inhabitant of that State in which he shall be chosen.

Representatives and direct Taxes shall be apportioned among the several States which may be included within this Union, according to their respective
Numbers, which shall be determined by adding to the whole Number of free Persons, including those bound to Service for a Term of Years, and excluding
not taxed, three fifths of all other Persons. The actual Enumeration shall be made within three Years after the first Meeting of the Congress of the United
and within every subsequent Term of ten Years, in such Manner as they shall by Law direct. The Number of Representatives shall not exceed one

Why is the government struggling with **Data Governance** ?

"Who makes the call?"



"What do you mean by *data dictionary* ?"

"Just give me all the data."

"This number doesn't seem right."

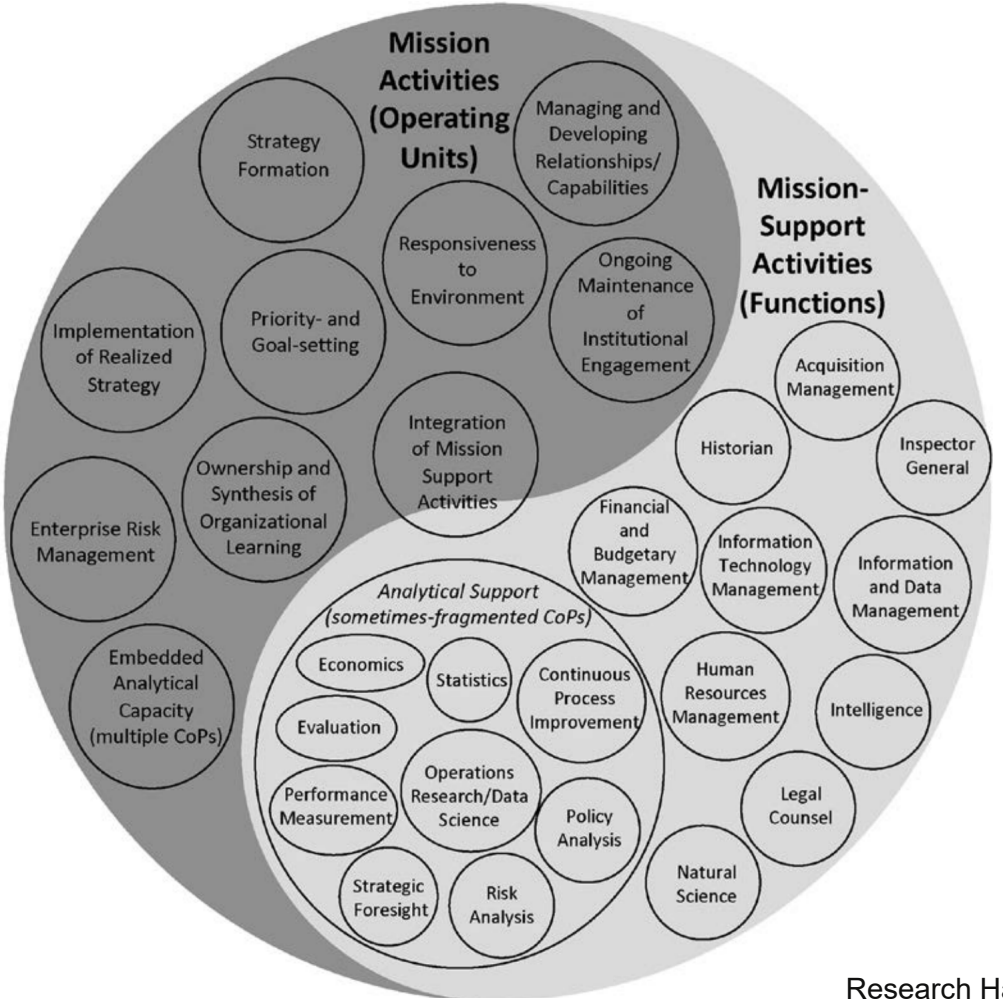
"I can't find what I need."

"Who can I ask about this metadata?"

"Is this Allowed?"

"Who can give me access?"

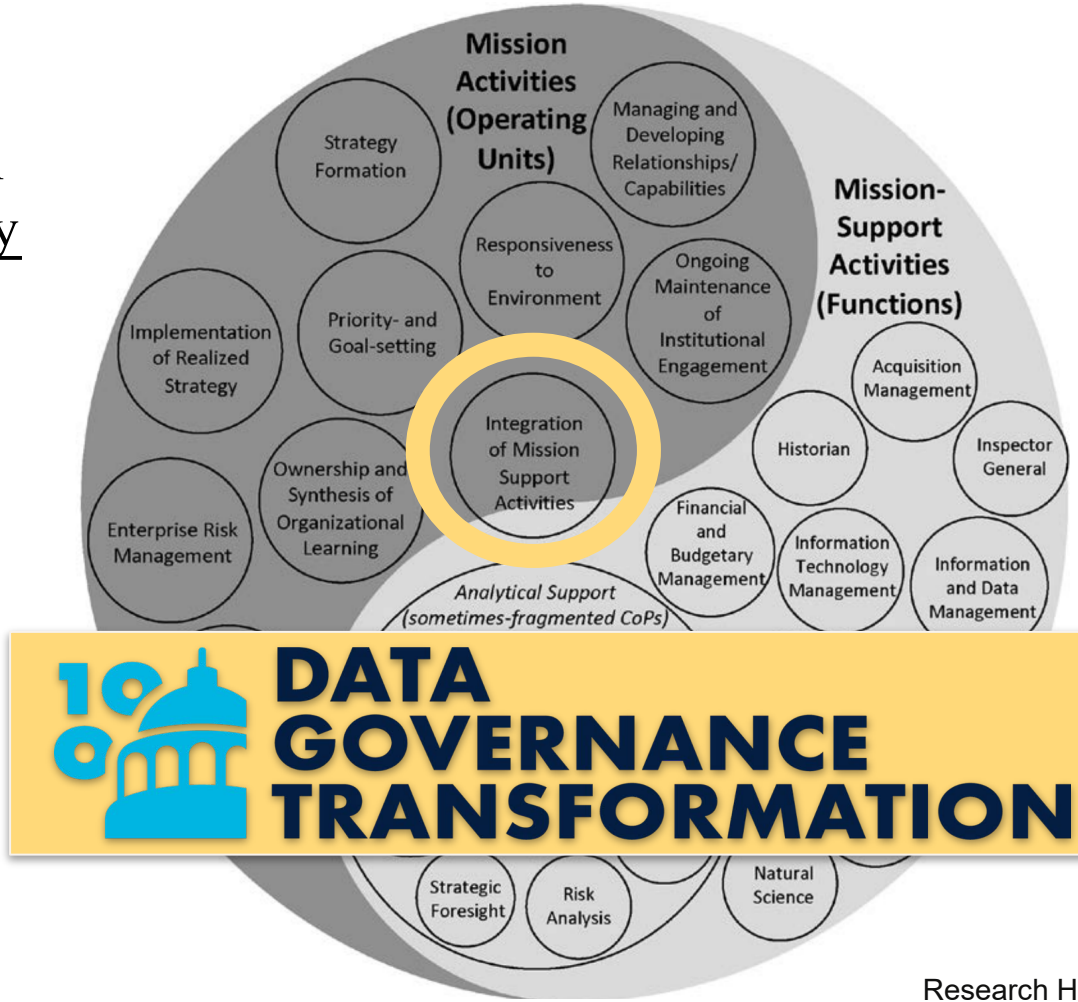
Mission Delivery



Mission Support

Research Handbook on Program Evaluation
Clinton T. Brass and Kathryn E. Newcomer

Mission Delivery



Mission Support

DATA GOVERNANCE TRANSFORMATION

Research Handbook on Program Evaluation
Clinton T. Brass and Kathryn E. Newcomer

Step 1: Consensus-Driven Creation of Data Governance Policy

Step 1: Consensus-Driven Creation of Data Governance Policy

Key Goals & Tenets

Step 1: Consensus-Driven Creation of Data Governance Policy

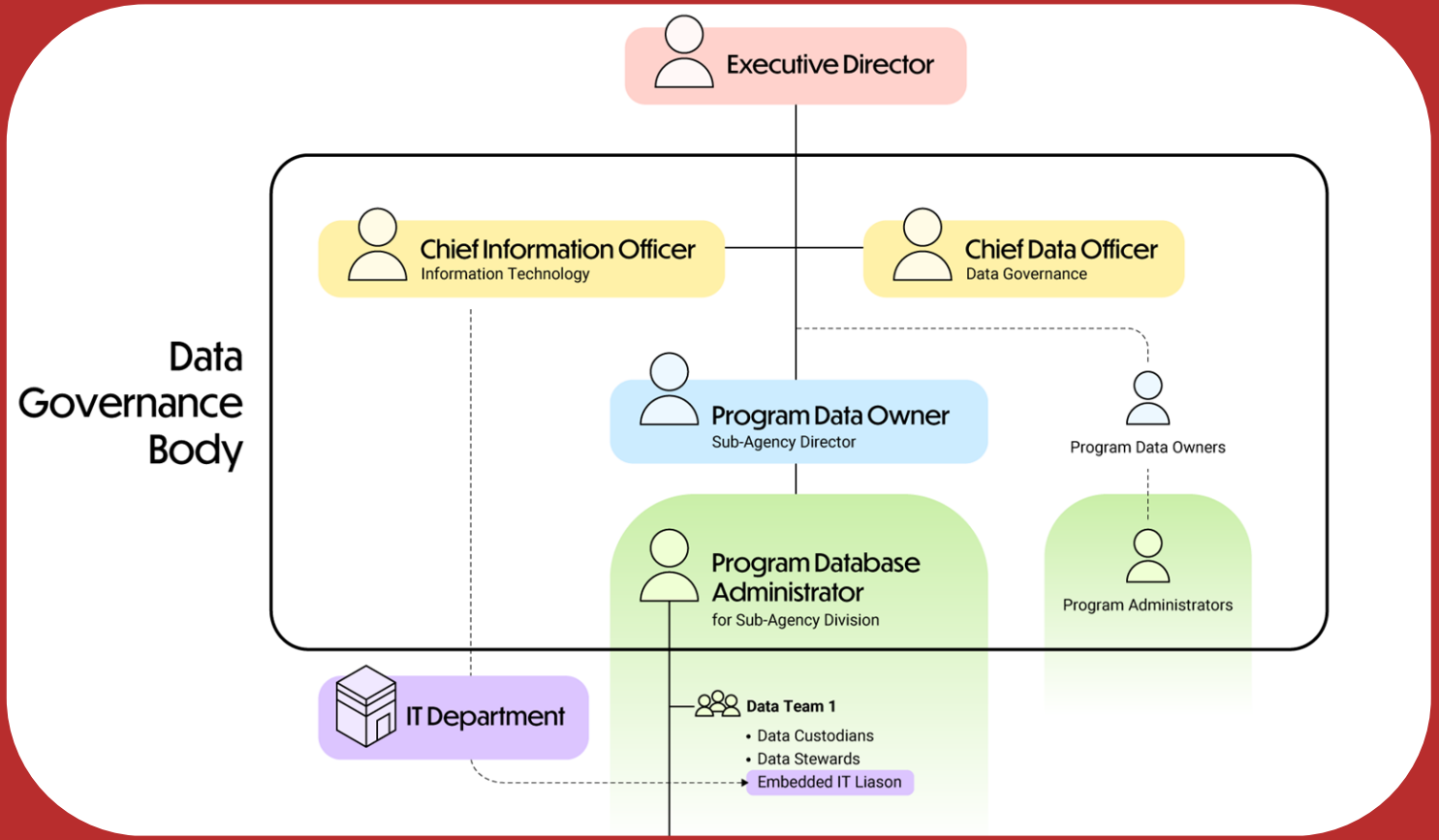
Key Goals & Tenets Rules of the Road

Step 1: Consensus-Driven Creation of Data Governance Policy

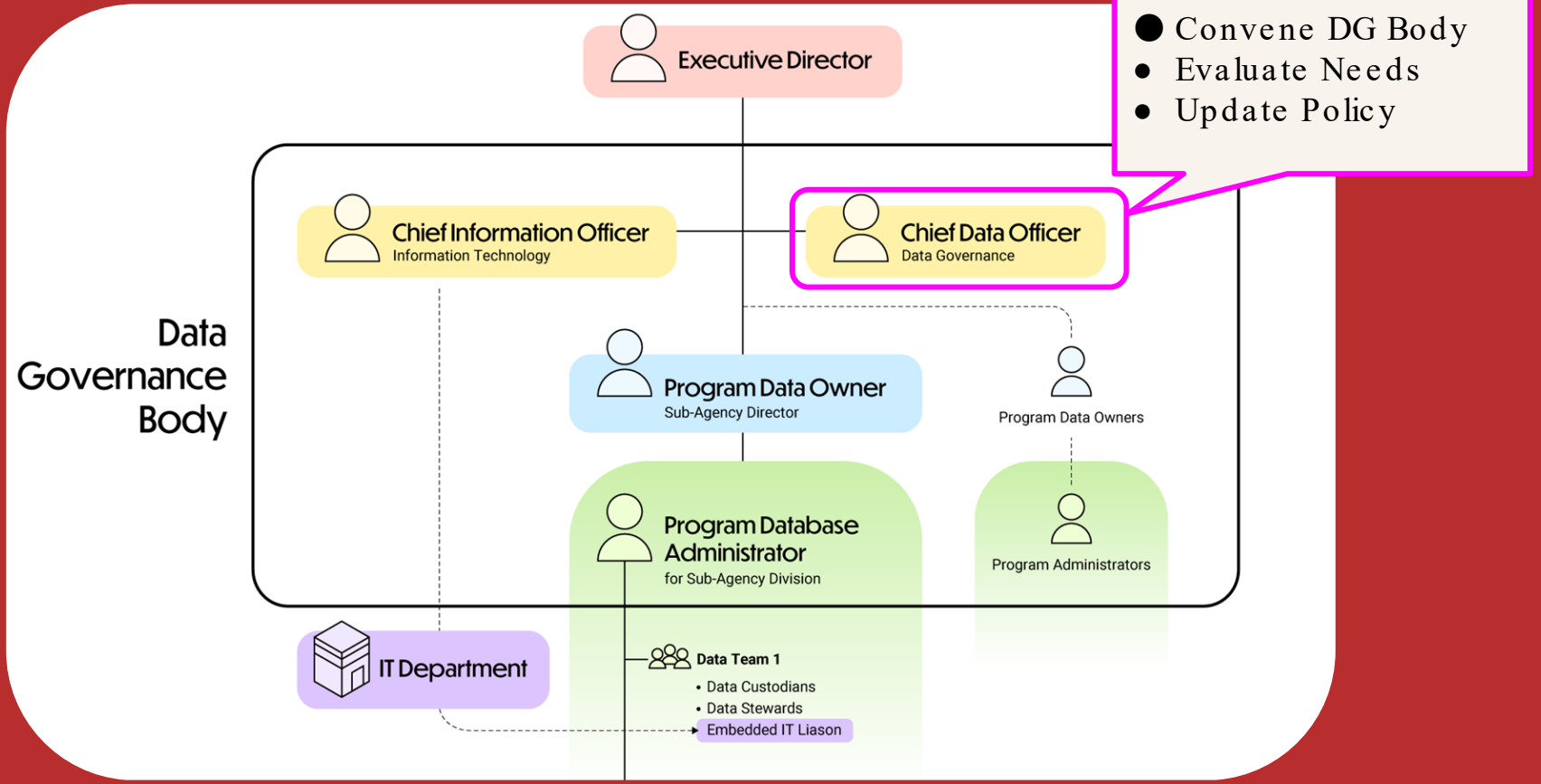
Key Goals & Tenets
Rules of the Road
Norms / Expectations

Step 2: Data Governance In Action

Step 2: Data Governance In Action

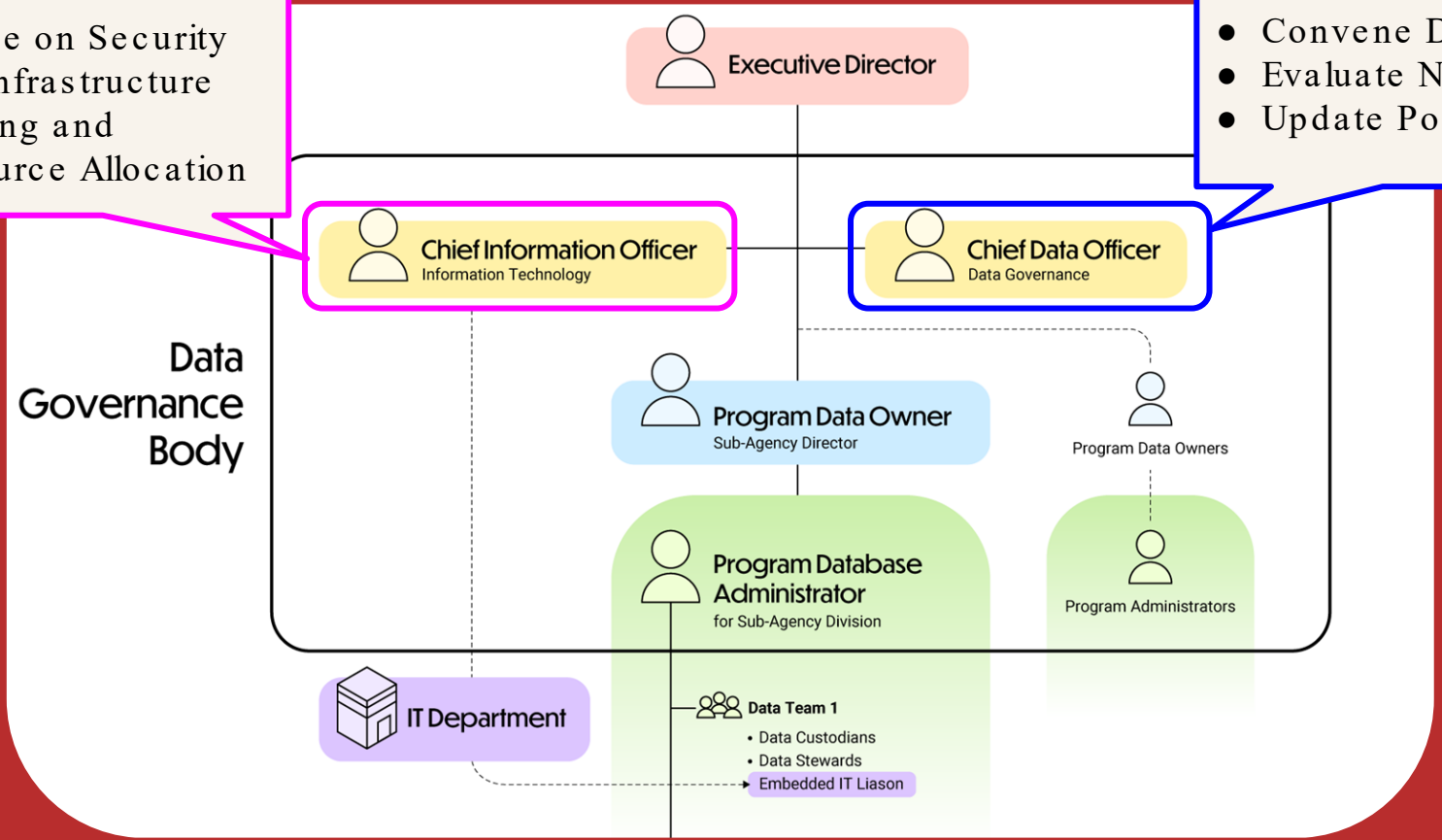


Step 2: Data Governance In Action



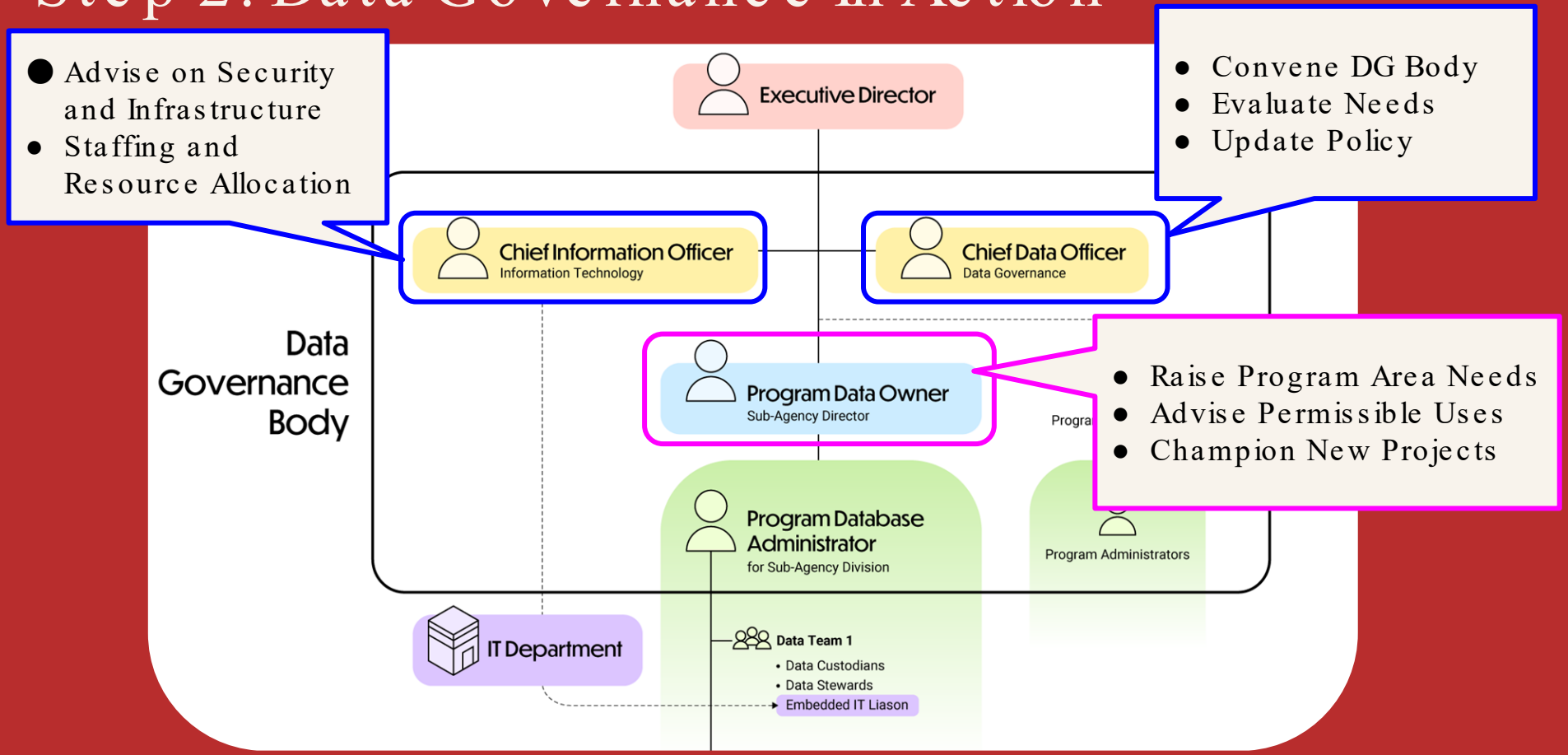
Step 2: Data Governance In Action

- Advise on Security and Infrastructure
- Staffing and Resource Allocation

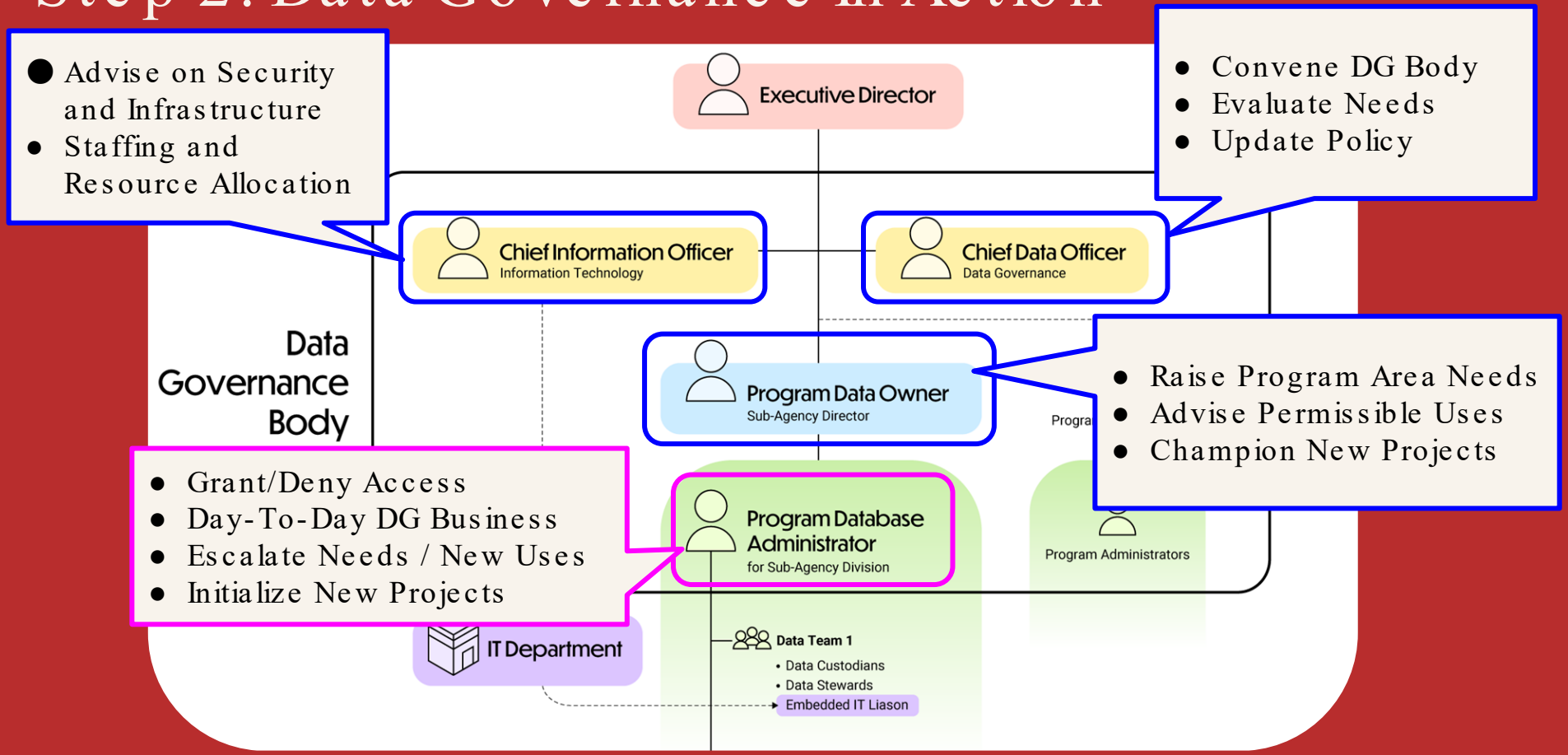


- Convene DG Body
- Evaluate Needs
- Update Policy

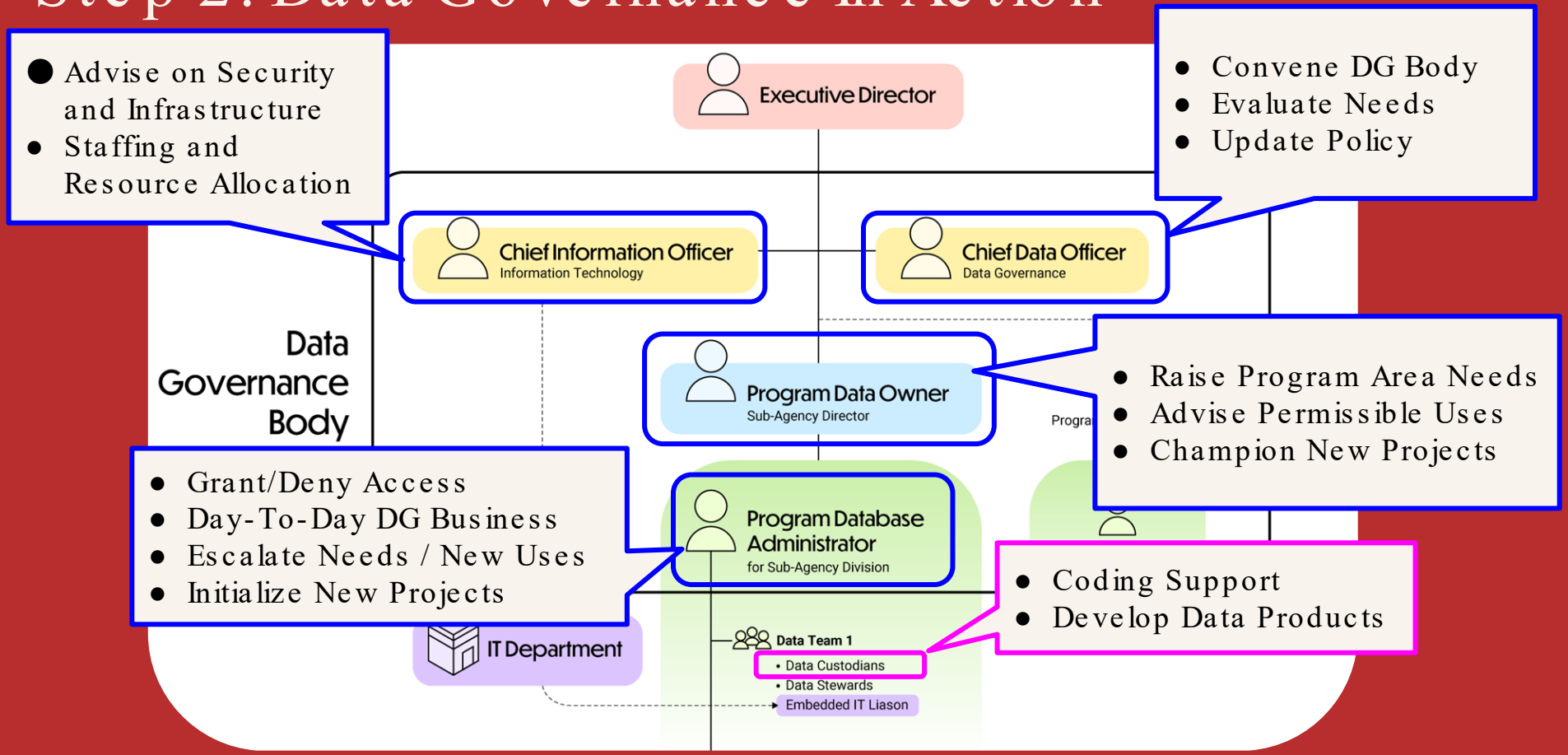
Step 2: Data Governance In Action



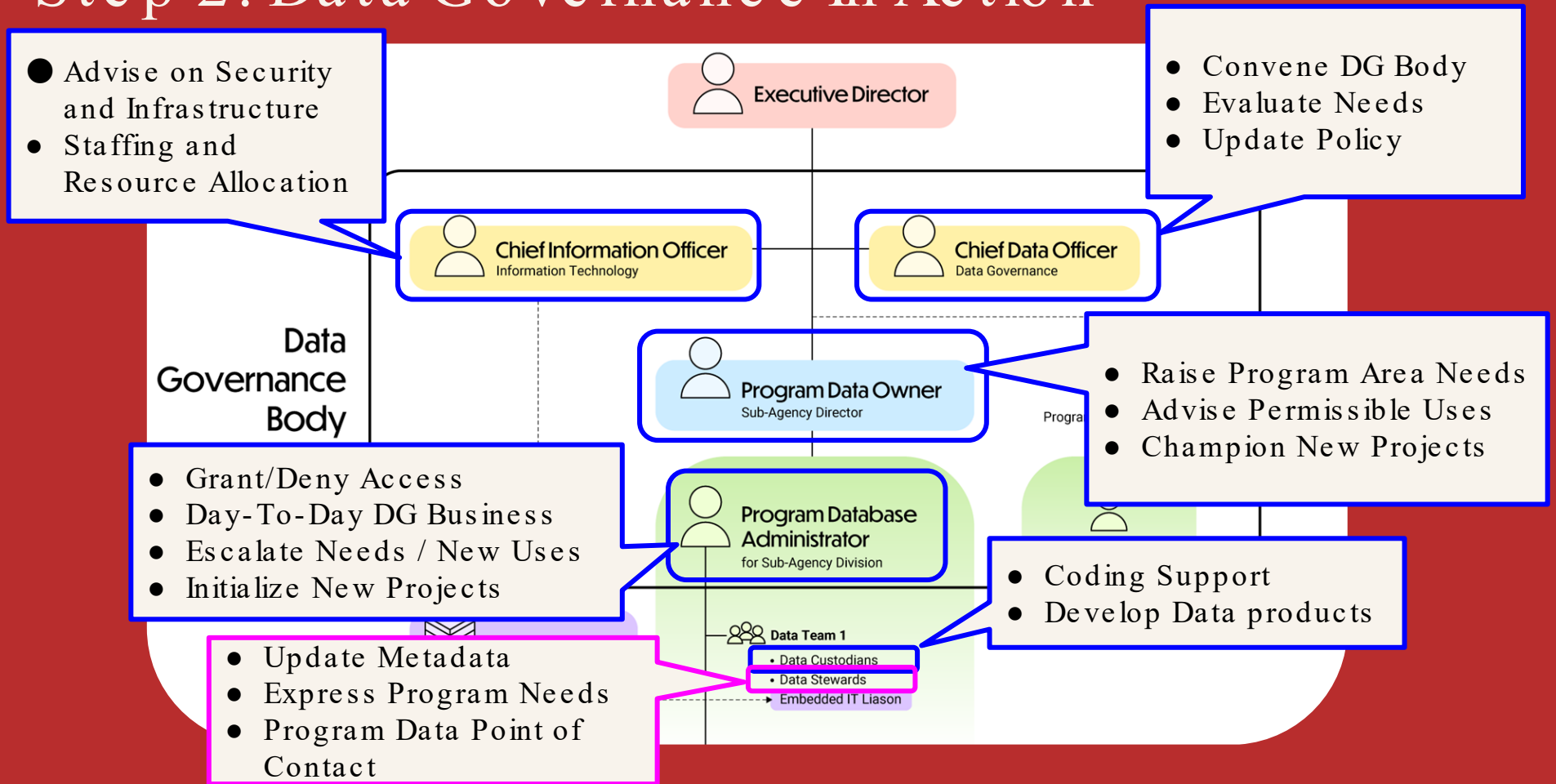
Step 2: Data Governance In Action



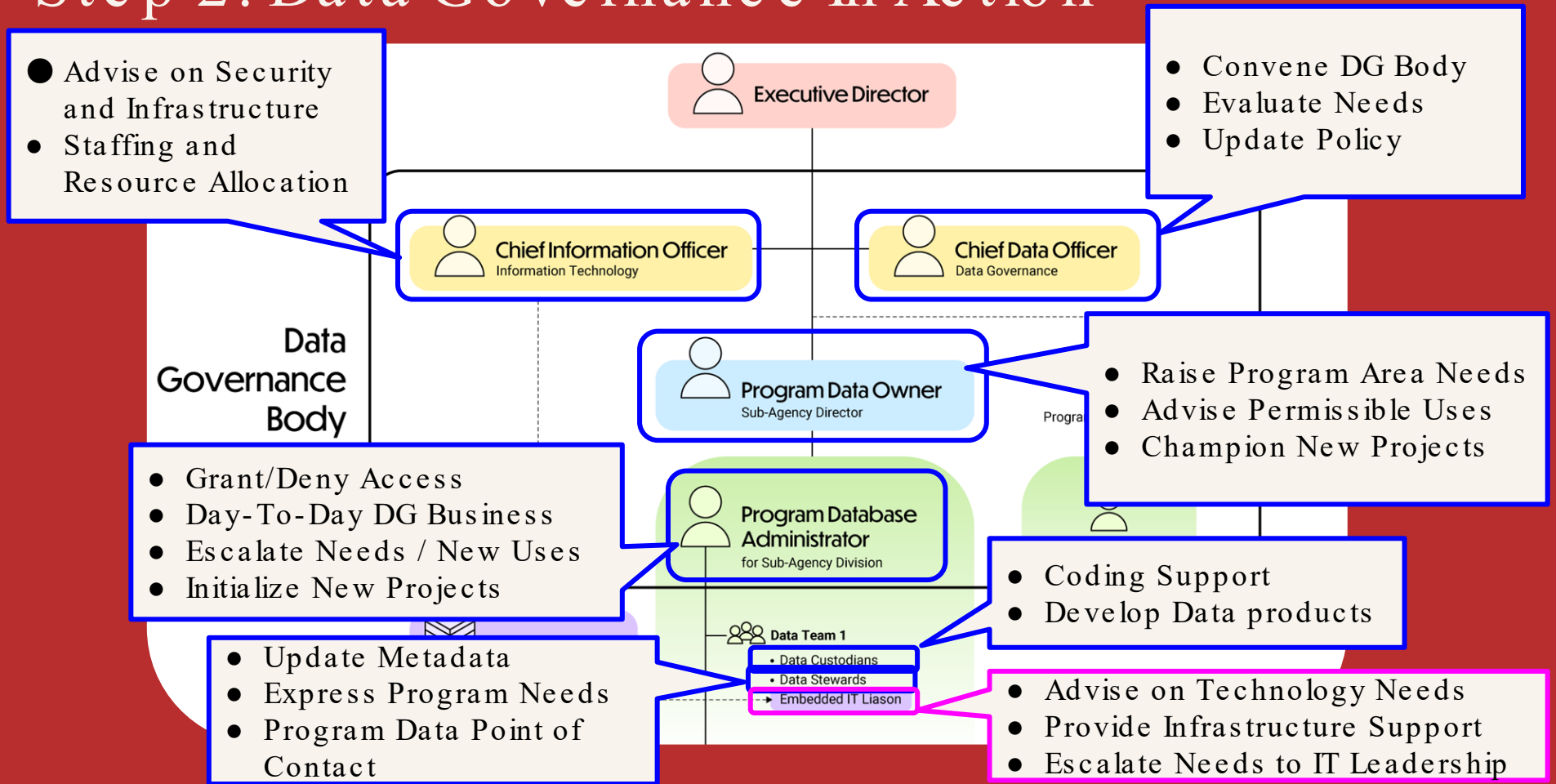
Step 2: Data Governance In Action

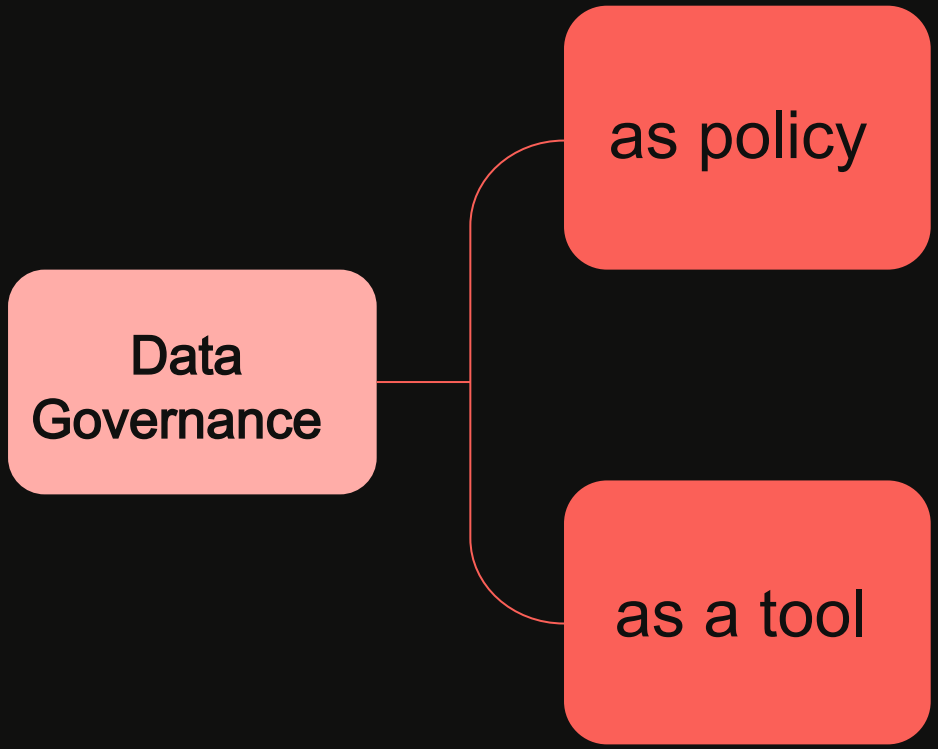


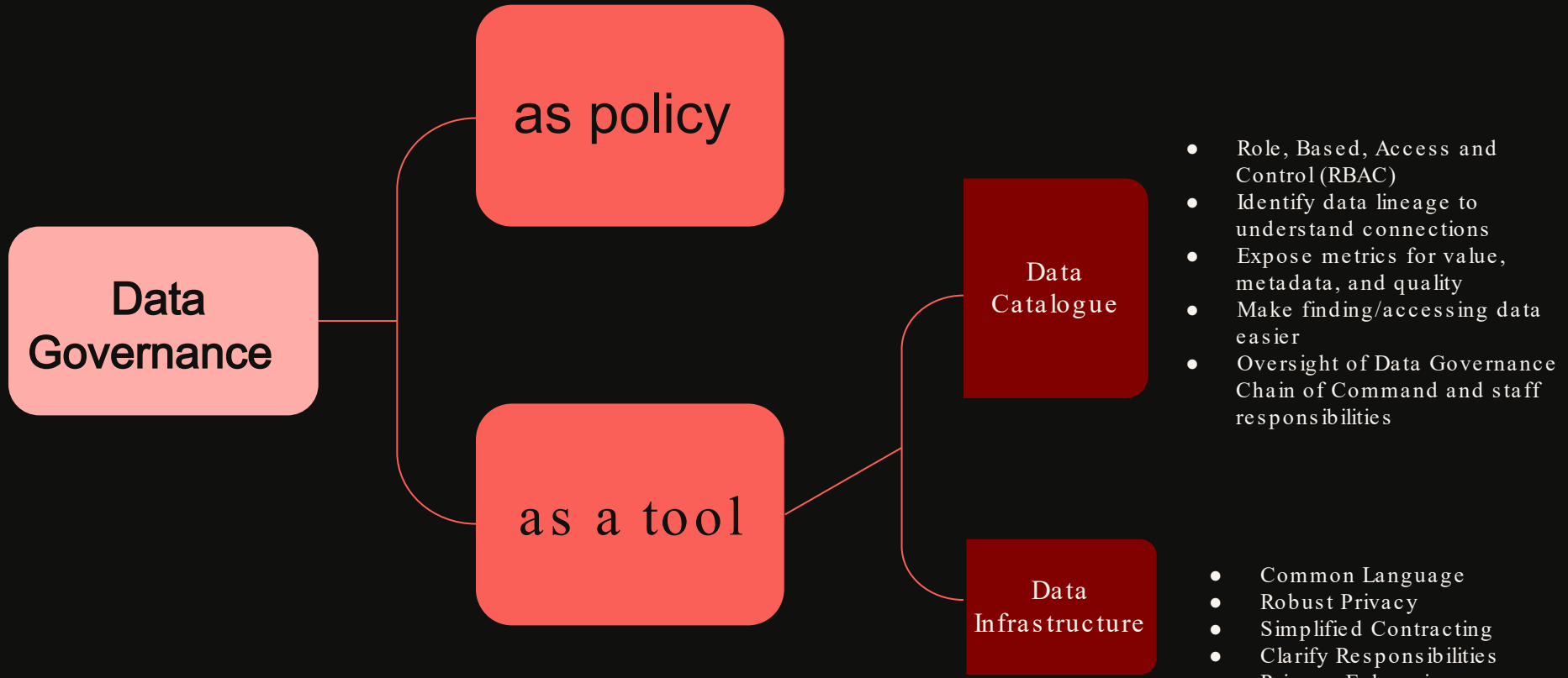
Step 2: Data Governance In Action



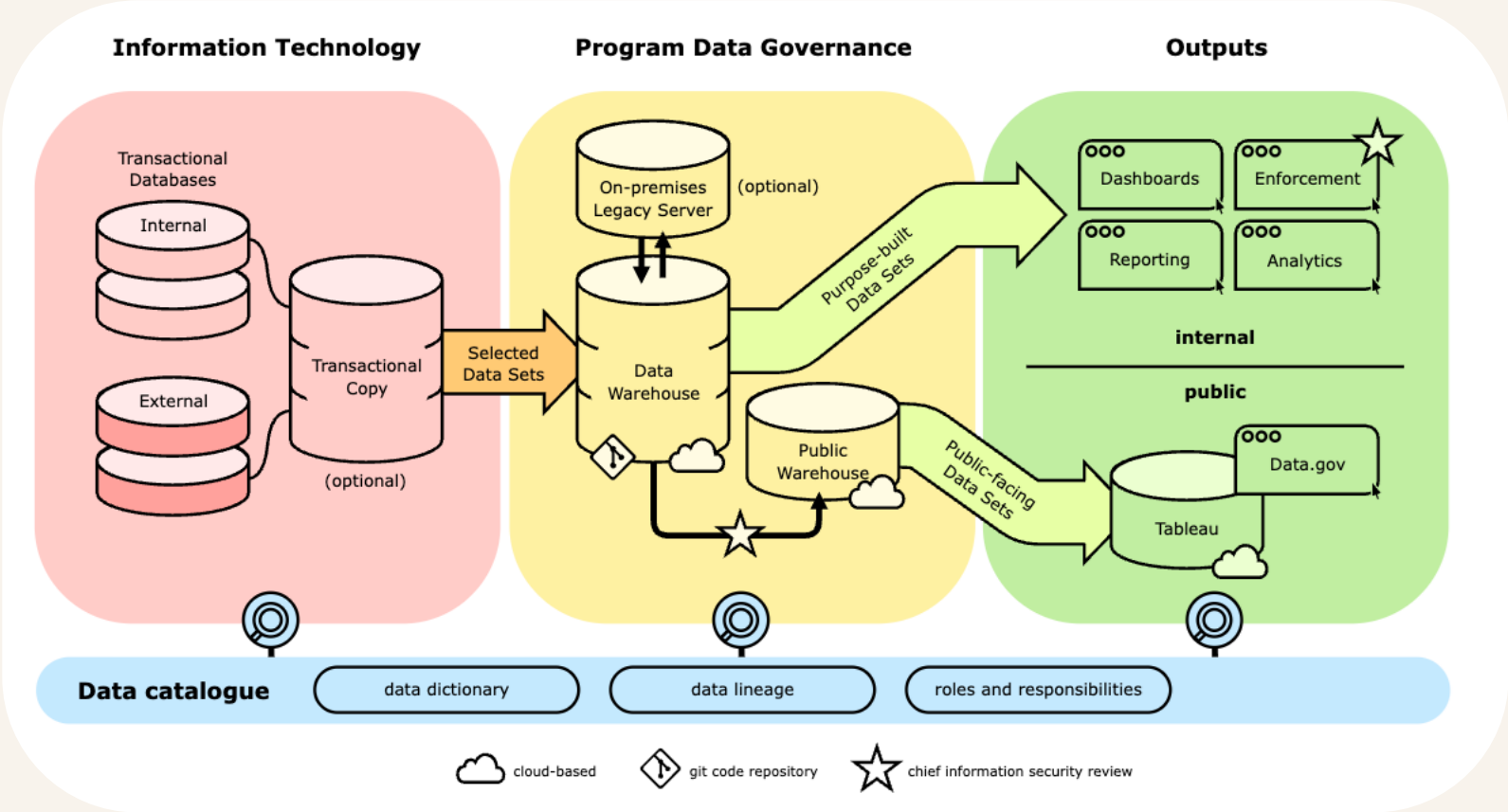
Step 2: Data Governance In Action




Data Governance
as policy
as a tool

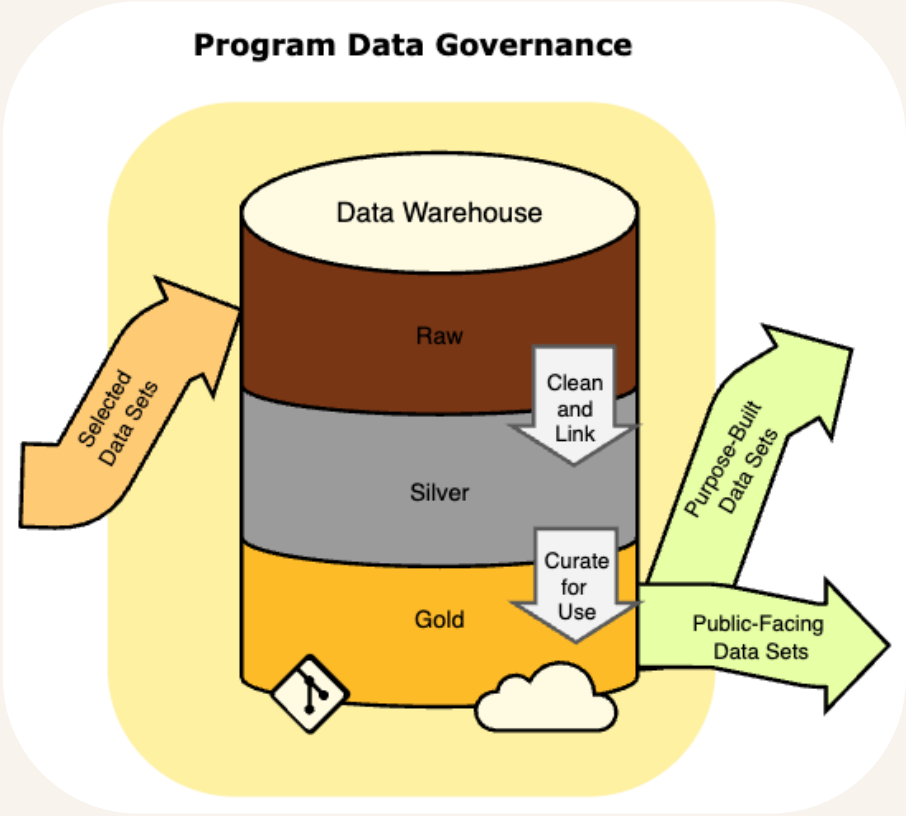

Data Governance
as policy
as a tool
Data Catalogue
Data Infrastructure
Role, Based, Access and Control (RBAC)
Identify data lineage to understand connections
Expose metrics for value, metadata, and quality
Make finding/accessing data easier
Oversight of Data Governance Chain of Command and staff responsibilities
Common Language
Robust Privacy
Simplified Contracting
Clarify Responsibilities
Privacy Enhancing Technologies

What does this new ecosystem look like?



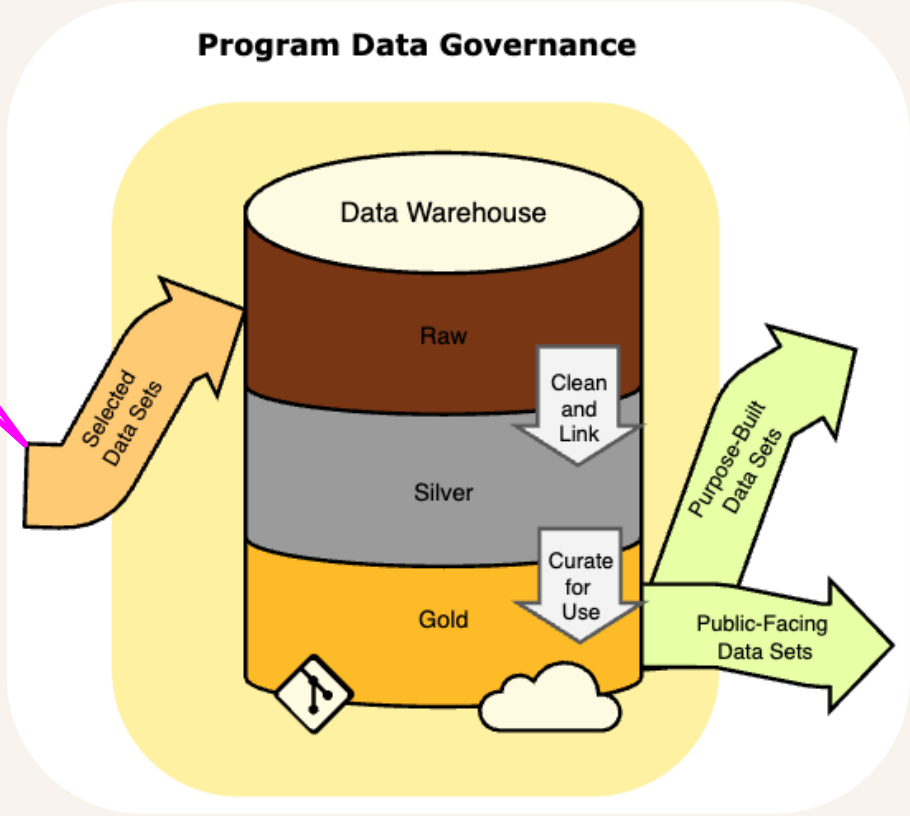
Let’s zoom in on the Program Data Governance…

Program Data Governance



Program Data Governance

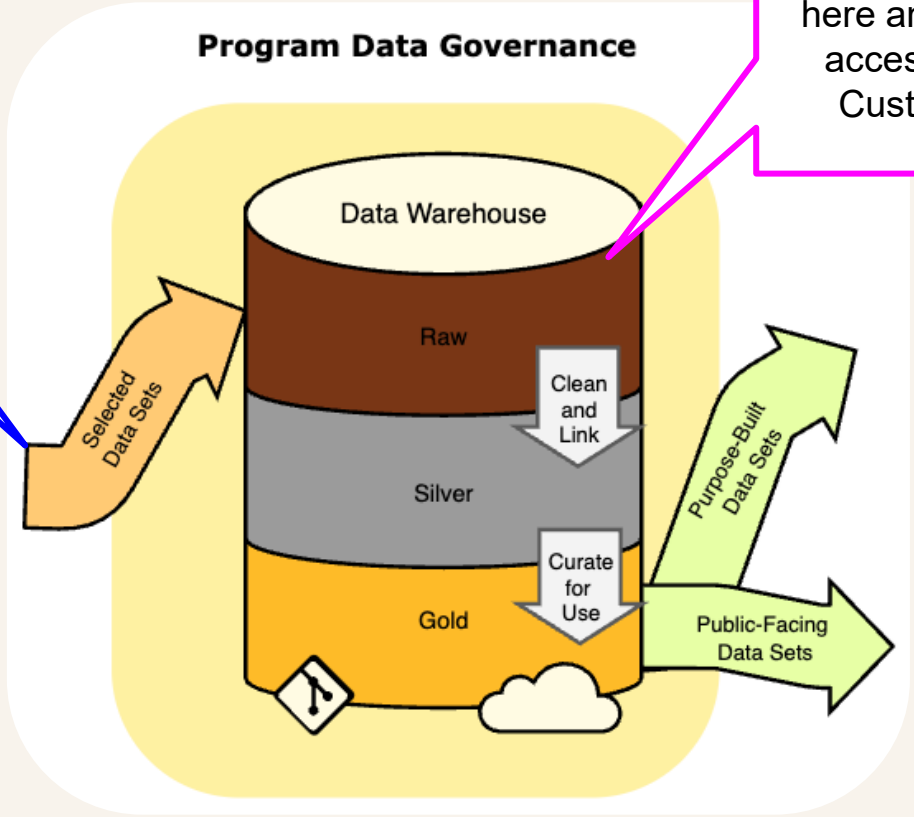
Data Custodians and Stewards work with IT Liaison to extract necessary data



Program Data Governance

Data Custodians and Stewards work with IT Liaison to extract necessary data

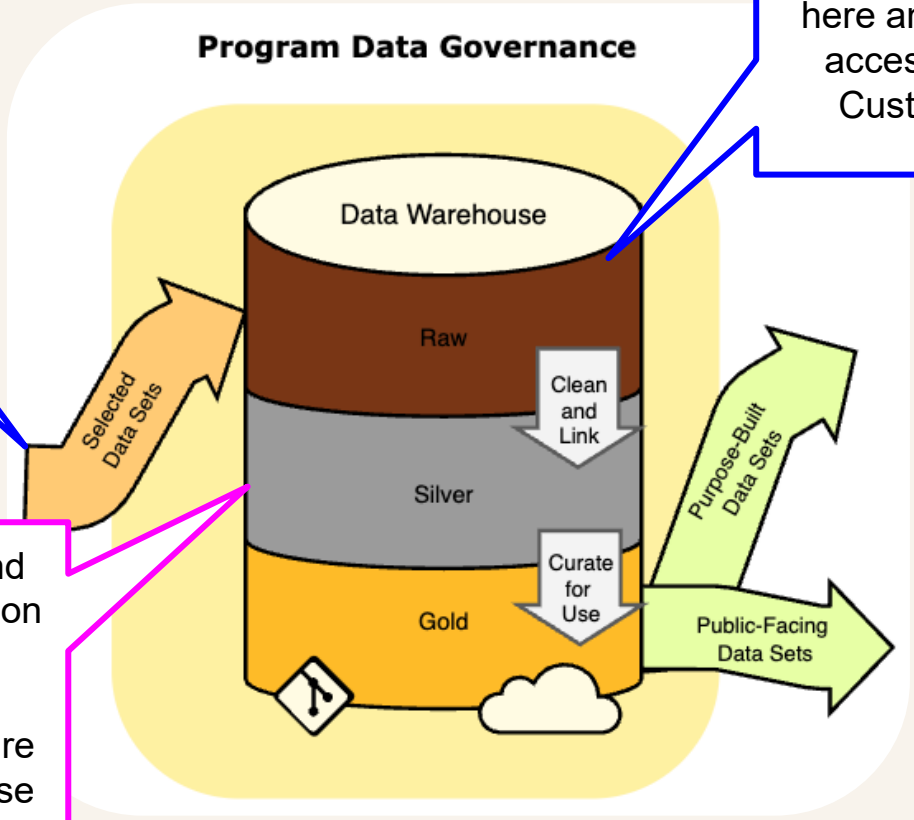
Raw data lives here and is only accessible to Custodians



Program Data Governance

Raw data lives here and is only accessible to Custodians

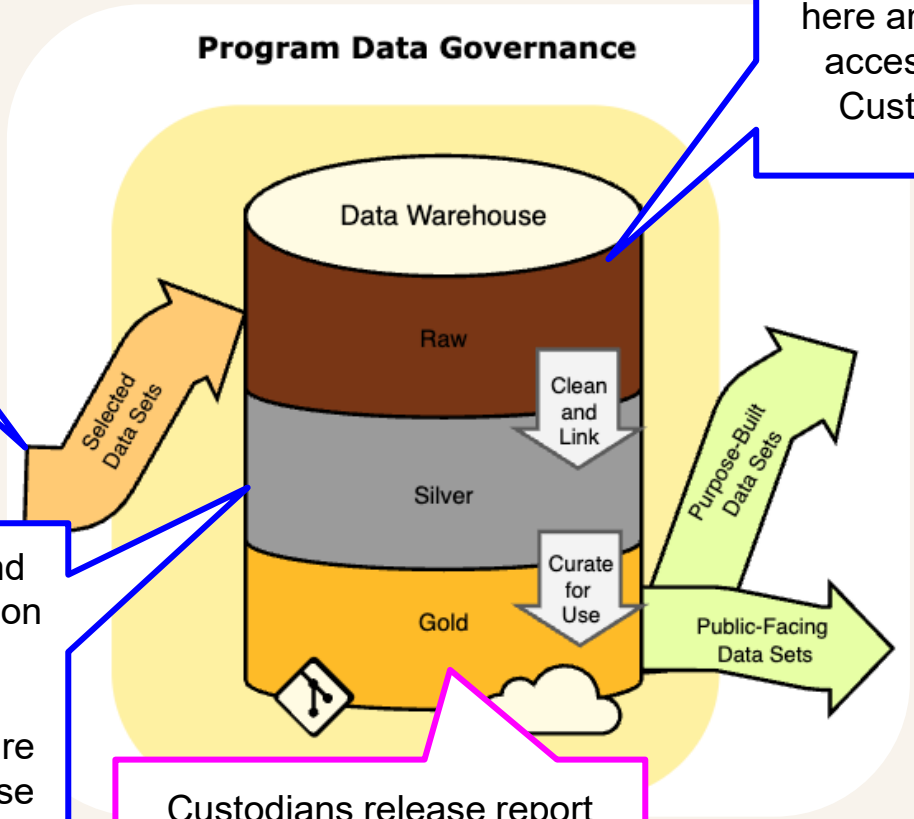
Data Custodians and Stewards work with IT Liaison to extract necessary data



Custodians clean, link, and transform data in preparation for use.

Stewards test products here but they aren't ready for use

Program Data Governance



Data Custodians and Stewards work with IT Liaison to extract necessary data

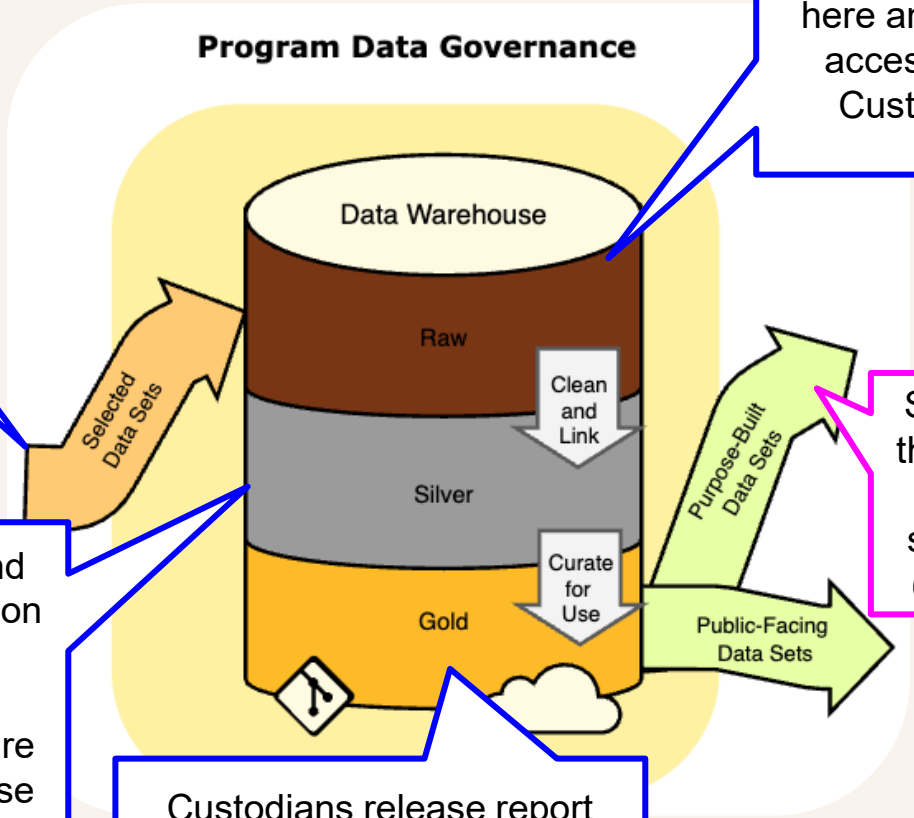
Raw data lives here and is only accessible to Custodians

Custodians clean, link, and transform data in preparation for use.

Stewards test products here but they aren't ready for use

Custodians release report ready products to appropriate Stewards for use after testing

Program Data Governance



Raw data lives here and is only accessible to Custodians

Data Custodians and Stewards work with IT Liaison to extract necessary data

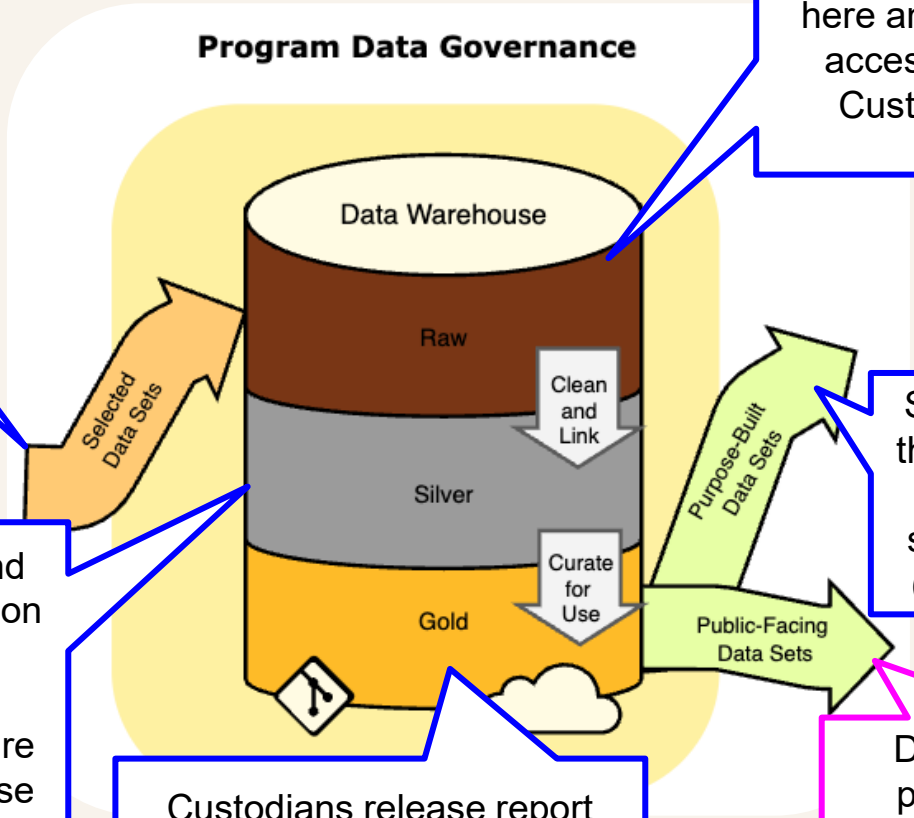
Stewards provide data to their Program teams in an easy to use form with support from Custodians (E.g. PowerBI, Tableau)

Custodians clean, link, and transform data in preparation for use.

Stewards test products here but they aren't ready for use

Custodians release report ready products to appropriate Stewards for use after testing

Program Data Governance



Data Custodians and Stewards work with IT Liaison to extract necessary data

Raw data lives here and is only accessible to Custodians

Stewards provide data to their Program teams in an easy to use form with support from Custodians (E.g. PowerBI, Tableau)

Custodians clean, link, and transform data in preparation for use.

Stewards test products here but they aren't ready for use

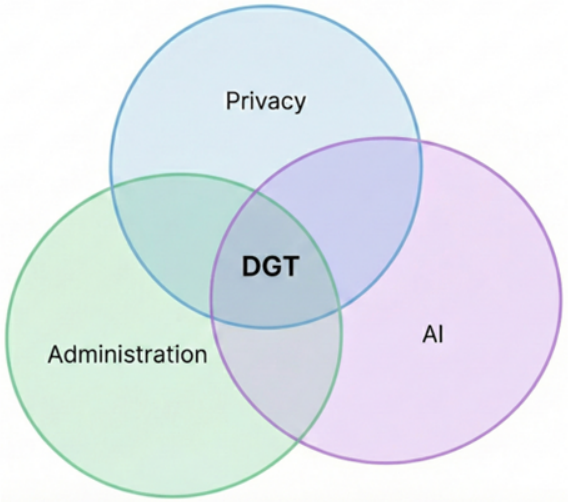
Custodians release report ready products to appropriate Stewards for use after testing

Data is released to the public after Disclosure Review. Privacy Enhancing Technology is used as needed

So how do we meet this moment?

"Data Governance" covers the tangle of data, systems, policy and people we need to navigate to select, confirm and deploy the solution. Data Governance Transformation is how organizations make sure they're **prepared** to successfully execute data projects.

Data **Governance Transformation** is the building block for any data project



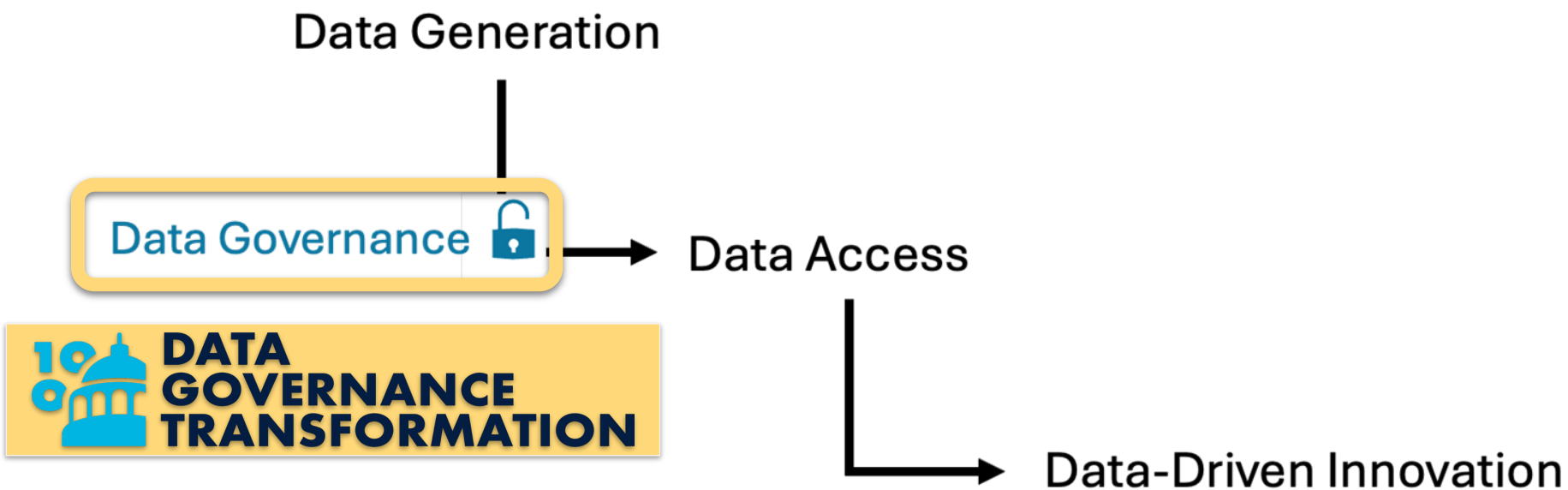
Privacy problem specification tasks must be done for **all** participating partners at the ***project outset***. All specification tasks are completed by members of the Data Governance chain of command of the participating partners.

Data Project Specification Checklist:

- **Data Categorization:** What is the sensitivity level of information as determined by agency policy and terms specified during collection.
- **RBAC:** Who is allowed to see/manipulate/grant access for any given data that will be implicated. This may include upstream data objects.
- **Data Architecture:** What data standards/validations/quality assurance/Primary Keys are implemented in the implicated databases.
- **IT Infrastructure:** Where is the data stored and what will need to be connected together to allow Authenticated users to do the required work.
- **Privacy Policy Requirements:** Which Privacy Policy regulations will be implicated for these data, their categorization, permitted uses, and the identified users.
- **Chain of Command:** Who is allowed to make decisions about allowable uses. Who will curate the data. Which stewards will be required to validate proper data usage.
- **Goal:** What is the expected outcome for all partners and how will this goal fulfill obligations/requirements/needs for them all (e.g. Permissible use for public good).
- **Transparency:** How can all participants (including people whose data is being used) keep up to date on project progress and outcomes to ensure proper usage and re-use of final data product.

Unlocking Data Access

Translating Data into Public Good



Next Steps

- Developing demonstration of DGT Data Catalog using Civil Justice Data Commons Data
- Implementing Data Governance Strategies at the Small Business Association
- Engaging with NIST to draft Data Governance Framework Standards
- Identify and partner with additional Federal Agencies ready to mature their Data Governance policies

Big Questions

Mailing List Sign - Up QR Code

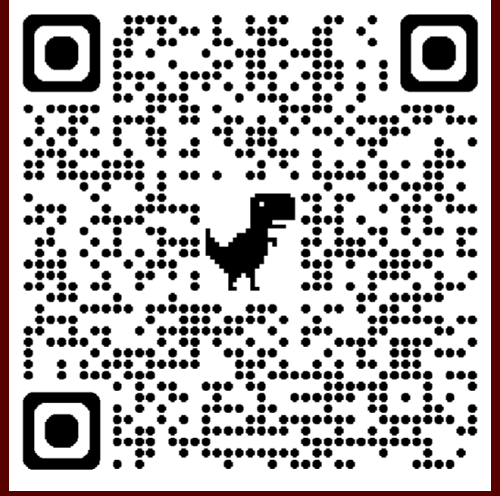
How can we make this structure more generalized so that it can be applied to any level of government?

What terms are opaque or might be misinterpreted by communities from different sectors?

How can we work with IT leaders to ensure this division of effort is fair and appropriate given the policy environment?

What roadblocks to implementation do you see?

DATA GOVERNANCE TRANSFORMATION



Thank you!

Jake.Pasner@georgetown.edu

DATA GOVERNANCE TRANSFORMATION