Western

Western Institutional Data Governance

Distribution Slides

October 2023

https://data.uwo.ca

Data Governance Framework

What Are Institutional Data?

Institutional Data are informational elements created, captured, curated, and managed within the inclusive context of Western's administrative and operational processes. Such datasets are typically arrayed within structured databases but can also be found as unstructured in nature.

Institutional data exists within the institution's enterprise resource platform (ERP), student information system (SIS), learning management system (LMS), customer relationship management (CRMs) systems, and other data environments that fall under the definition of institutional *systems of record* (SoR). Institutional data may also reside within divisional or departmental contexts as *line of business* systems (LoB), or may be housed within granular *local* solutions such as forms, Excel spreadsheets, Access databases, etc.

Western's data systems may also contain sensitive information not intended for institutional analytics, but they are still considered part of Western data governance in terms of defining purpose and roles and responsibilities. Such data repositories would include metadata and certain forms of transactional data.

Data Governance Framework

What Is Data Governance?

Data governance is the process through which consistent and clear guidance around defining, creating, managing, accessing, and using Western's institutional data are organized and facilitated. Data governance is a process used to promote and support the responsible use of high-quality institutional data, to facilitate informed and insightful use of these data, and to increase their value to the Western community.

Data governance includes attention to relevant risks, responsibilities, and legal obligations that are important and/or required for the institution. Data governance includes the definition and specification of decision rights and an accountability framework to encourage empowered and responsible behaviour in the valuation, creation, storage, use, archival, and destruction of data. Data governance is largely concerned with practices and processes which are used to ensure formal management of an organization’s data assets, while attending to data leadership, stewardship, quality, sharing and access, security and privacy, and reliability.

Western’s data governance framework does not pertain to data collected for scholarly research.

Data Governance Framework – Guideposts

Visibility

Sets of processes associated with strengthening the processes and elevating awareness related to Western's institutional data assets, the ways in which data assets can and should be used, and the information lifecycles involved. Ensuring that there are common data definitions and that those definitions are made available across platforms is essential to enabling informed data-driven decision-making. This pillar will be used to make decisions on what those definitions are and how we technically support the requirements of those definitions.

Accountability

Data leadership and stewardship are essential parts of any data governance program. Those responsible for data leadership are accountable for the integrity and quality of our data. This pillar will be used to determine who in the institution has the authority to make decisions regarding access, priorities, and data usage standards, and under what conditions those decisions can be made.

Reliability

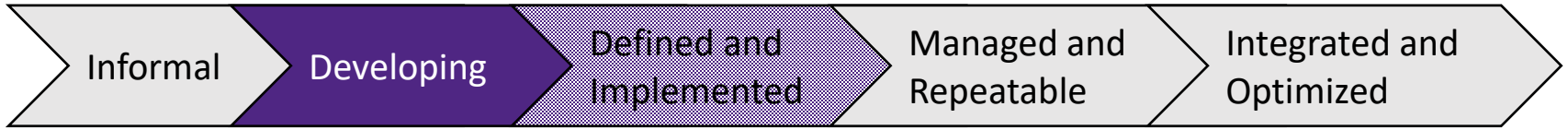
Processes and metrics to assist in the definitions and measurements of data quality, the attendant availability of institutional data, and the usefulness of said data within a variety of strategic and tactical contexts. This pillar is concerned with optimizing data and related data analytics infrastructure and services.

Usability

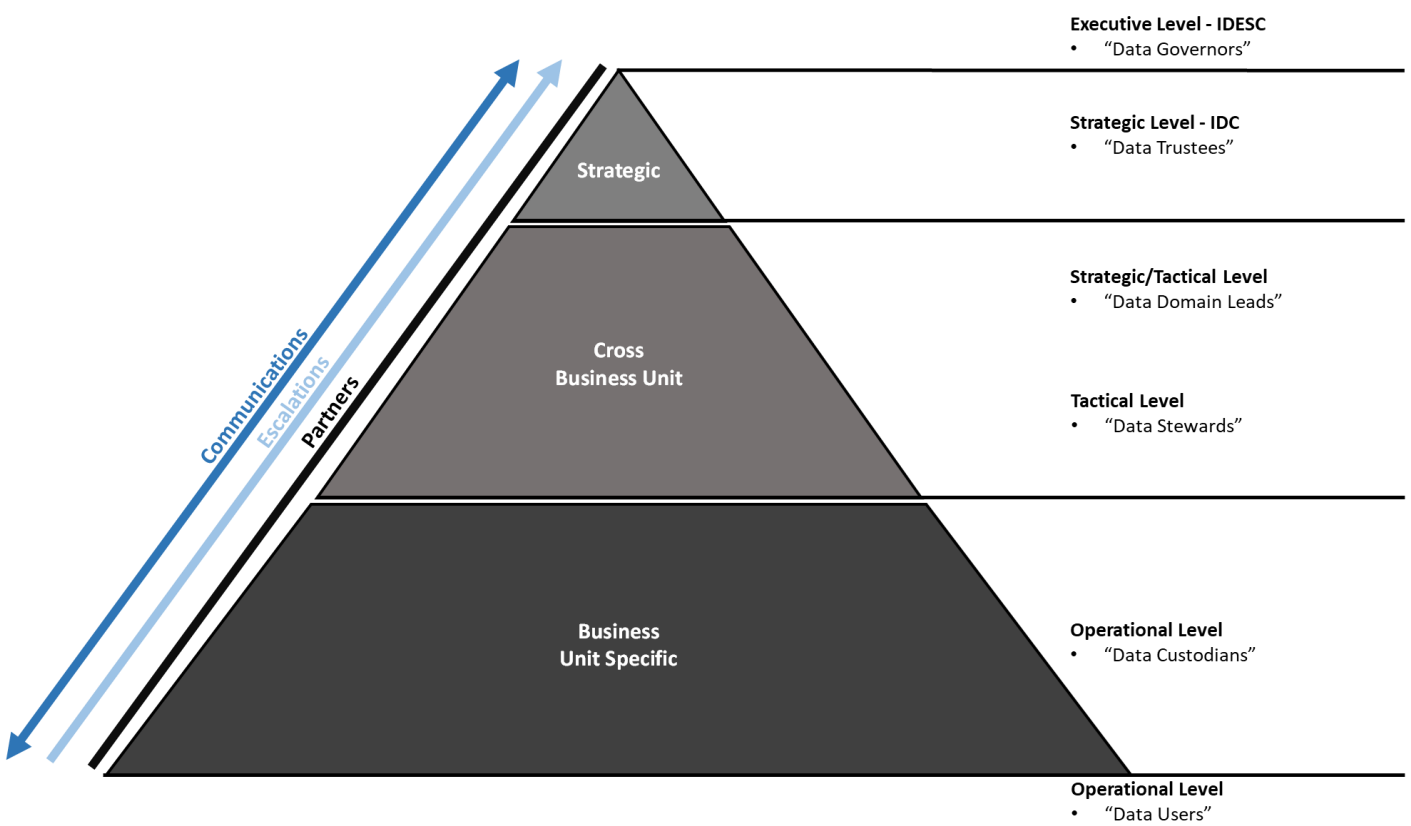
Working through the defining, creating, storing, sharing, archiving, and destruction phases of the information lifecycle, clear documentation that outlines data timeliness, relevancy, lifespan, and usefulness will be measured via this pillar. This pillar will be used to empower our data community through data literacy, manage access to relevant and appropriate datasets, and ensure that data are used in compliance with regulatory and institutional contexts, with privacy, security, and sensitivity embedded in our processes.

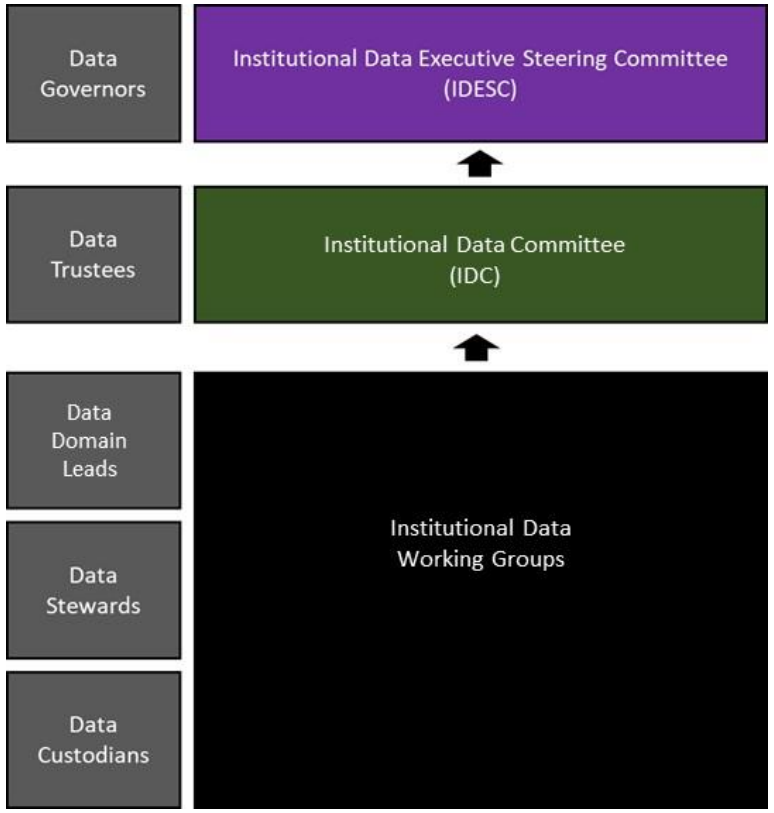
What Does Data Governance Maturity Look Like?

- Reliable Data Sources
- Clear Data Roles
- Authoritative Data Definitions
- Data Conforming Processes
- Central Data Warehouse
- Institutional Data Policy
- Impactful Sharing Processes
- Institutional Dashboard Portal
- Relevant Historical Contexts
- Robust Analytics Programs



Data Governance Relationship Diagrams

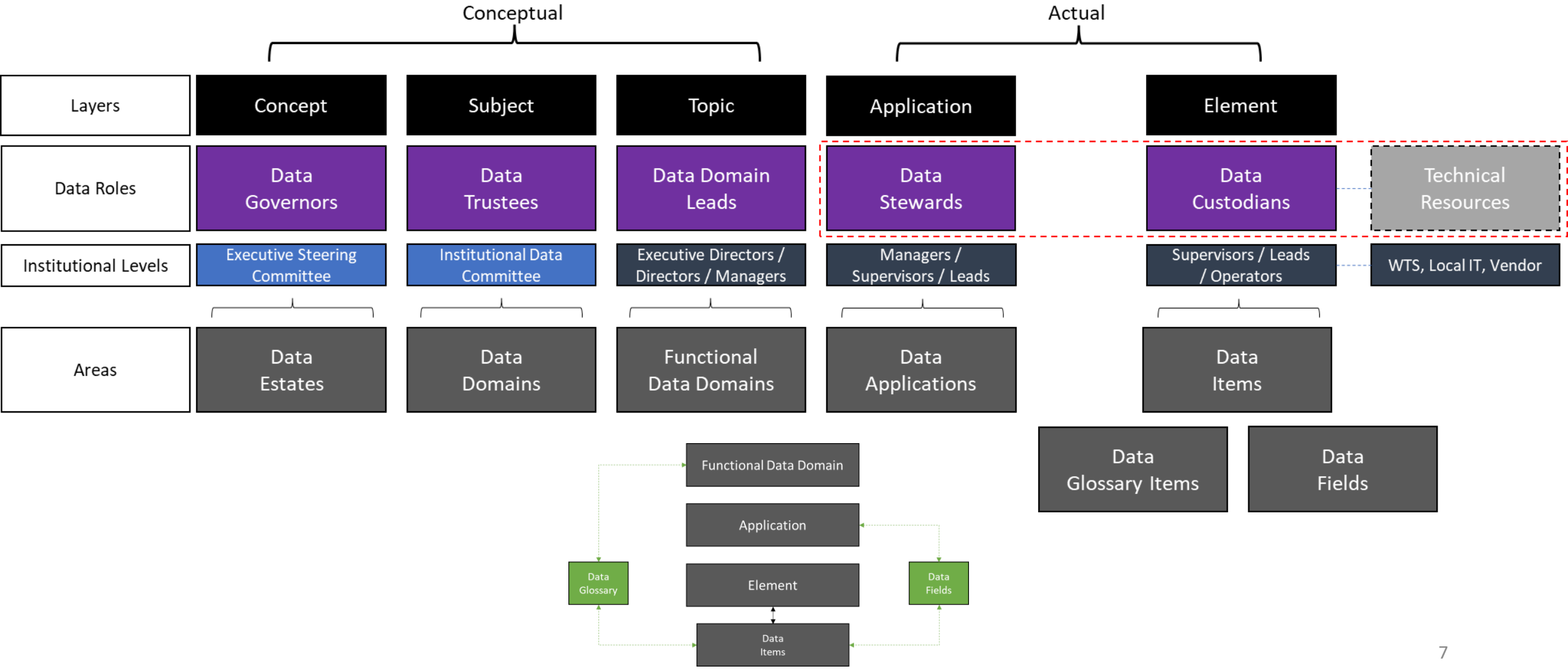




Strategic oversight is provided by the Executive Steering Committee and the Institutional Data Committee acts as the primary decision-making body for cross-university data ecosystem contexts.

Working Groups will be established as either "standing" or "ad hoc", depending on the need. the Working Group on EDI Data and a Working Group on Reporting and Analysis are good examples of standing Working Groups.

Data Governance Map



Data Lineage Hierarchy Descriptions

Data Estates (*Concept*)	SHARED ACCOUNTABILITY

Highest level of data presence at the institution.

"People", "Research", "Space", "Academic", "Services"*

Data Domains (*Subject*)	SHARED ACCOUNTABILITIES

Highest level of logical data groupings for the purposes of brokering the relationship between data estate and functional data domains.

Examples: "Student", "Employee", "Facilities", "Planning", "Accounting", "Budgeting", "Information Technology…"

Functional Data Domains (*Topic*)	PARTNERED ACCOUNTABILITIES

Exists at the level of divisible relevance within the institution and is the location for high-level data leadership attribution.

Examples: "Undergraduate", "Graduate", "Buildings", "Grounds", "Infrastructure", "Cyber Security…"

Data Systems (*Application*)	MULTIPLE RESPONSIBILITIES

Refers to the utility of data within the institution and usually software that function as systems of record (SoR), line of business (LoB), or local in orientation.

Examples: "PeopleSoft", "Salesforce, "Orbis Outcomes Plus", "OWL"…

Data Items (*Elements*)	MULTUPLE RESPONSIBILITIES

Examples: "EMPLID" (employee or student numbers), "First Name", "Last Name", "Username", "Email Address", "Session ID", "Course ID"…

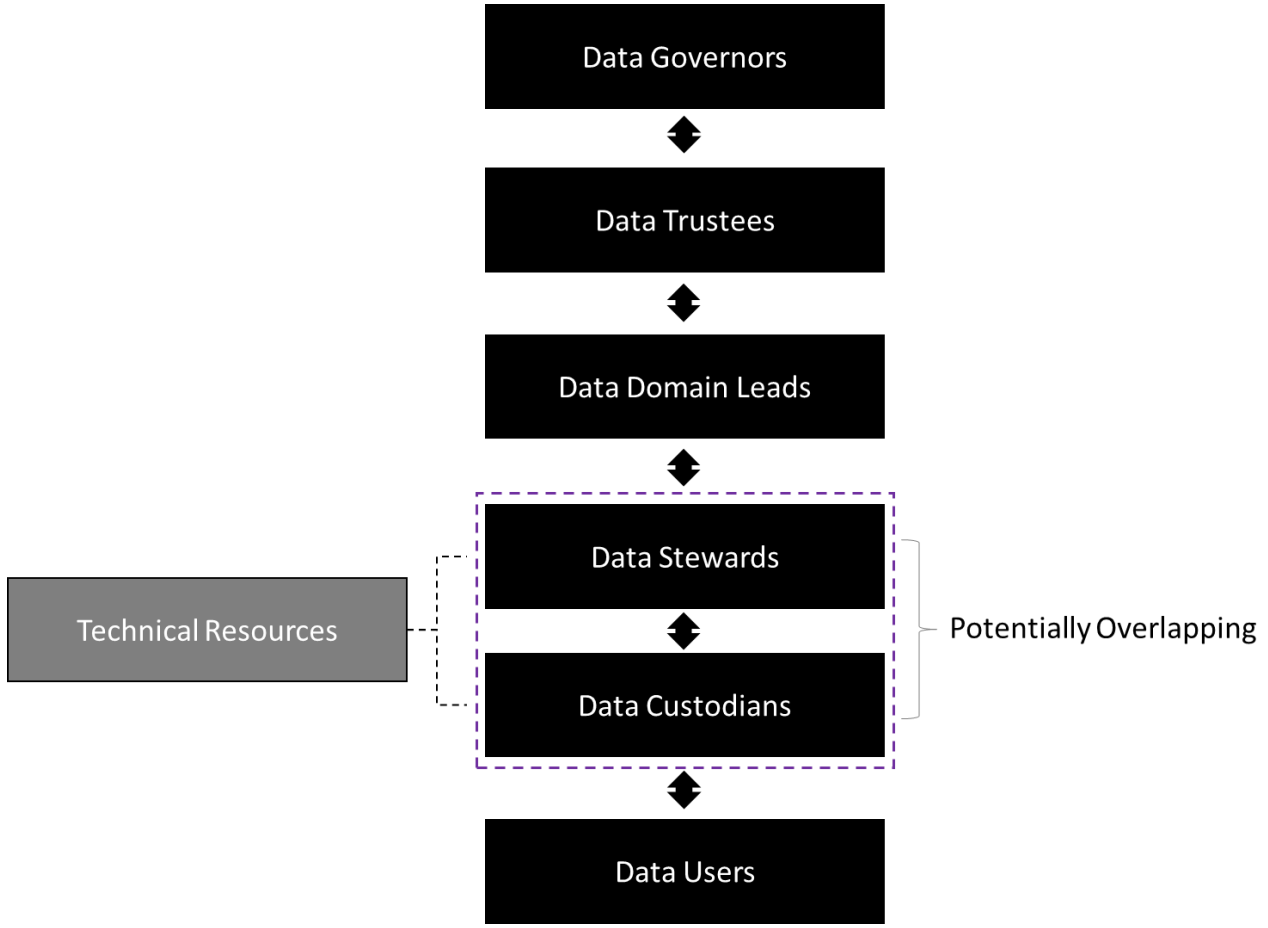
Data Lineage Examples:

CONCEPT > **SUBJECT** > **TOPIC** > **APPLICATION** > **ELEMENT**

PEOPLE > **STUDENT** > **UNDERGRADUATE** > **APPLICATION** > **EMPLID**

PEOPLE > **EMPLOYEE** > **STAFF** > **PEOPLESOFT** > **EMPLID**

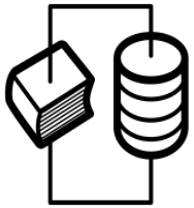
Data Roles and Accountabilities



The relationships in this diagram will be represented in our Data Asset Management Platform which will allow for strong approval governance for ongoing activities.

Formal structures at the Data Lead and Data Steward levels will provide a context for visibility of data activities across the institution.

Data Roles and Accountabilities



Data Governors

Data Governors are the executive functions of the institution comprised primarily of the Vice-Presidents and select leaders. Data Governors are accountable for the strategic leadership of the institution and are members of the Institutional Data Executive Steering Committee (IDESC). Data Governors are what we would consider as the Data Owners for all institutional data at Western.

Data Trustees

Data Trustees are senior institutional leaders with a strategic, cross-functional view of the over-arching data domains in play at Western. These roles promote data governance within the division's business and technical processes, resolve internal domain-oriented conflicts, look for linkages with other domains, and escalates to the IDESC when necessary. Data Trustees are accountable for ensuring the integrity, security, and use of Institutional Data within Western and for developing and recommending, and supporting and enforcing, relevant institutional policies, standards, and practices. Data Trustees appoint Data Leads for respective data domains. Data Trustees are members of the Institutional Data Committee (IDC).

Data Domain Leads

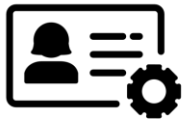
Data Domain Leads are the roles that are accountable for data presence at the functional data domain level. Responsibilities include formal approval of data definitions, data classifications, data sharing agreements, approval of data access requests, and determination of ongoing departmental strategy. A Data Domain Lead is also concerned with risk, appropriate access to data, and the presence of a solution’s data within the institution. These roles actively champion the ideation, implementation, and enforcement of data governance within their functional areas of responsibility.

*Note about Roles: Given Western’s complexity and historical processes, it is possible that multiple data roles can be assumed by singular institutional roles and people.

Data Roles and Accountabilities

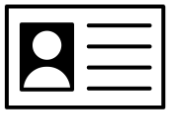
Data Stewards

Data Stewards are highest level of specific domain expertise in a given area (knowledgeable about data definition, lifecycle, history, sharing, and use) and are largely concerned with application of data within data systems. Formally responsible for data presence within the institution and effort would include the ability to amend data definitions, data classifications, and data sharing arrangements. A Data Steward is concerned with the meaning of data and the correct use of data within respective contexts. Data Stewards are accountable to Data Leads within the data governance model specifically. Data Stewards lead in defining, implementing, and enforcing data management processes and procedures within their application context.



Data Custodians

Data Custodians are roles identified as creators, curators, updaters, and sharers of institutional data within specified contexts. Formally responsibility for data reliability and trustworthiness and efforts include informing data definitions and data classifications. A Data Custodian manages the actual data in the form of processes, extract-transform-load (ETL) procedures, access control, backups, and documentation. The Data Custodian is responsible for documenting solution schema and data lineage. Data Custodians are system administrators or application operators who are responsible for the management of solutions that collect, manage and provide access to institutional data.



Data Users

Data Users are those whose roles require access to institutional data and where such data are used to inform reports, institutional and unit-level decision-making, and/or exploratory analysis; Data Users include data producers, reporters, consumers, and brokers within the context of day-to-day operations.

*Note about Roles: Given Western's complexity and historical processes, it is possible that multiple data roles can be assumed by singular institutional roles and people.

Data Catalog, Dictionary, Requests, Sharing, Deep Dives

Data Catalog and Data Glossary (Dictionary)

A data asset catalog is a repository that will assist in expanding institutional knowledge about the kind and nature of data systems and elements in use. The data catalog will be mapped to the entity framework so that roll up reporting can be achieved at data domain and functional data domain levels. It is proposed that data assets be registered into the catalog and processes such as TRAC and auditing can assist in ensuring compliance. Likewise, a data glossary (or dictionary) will define for the institution the various terms and vocabulary used to describe processes and/or data elements themselves.

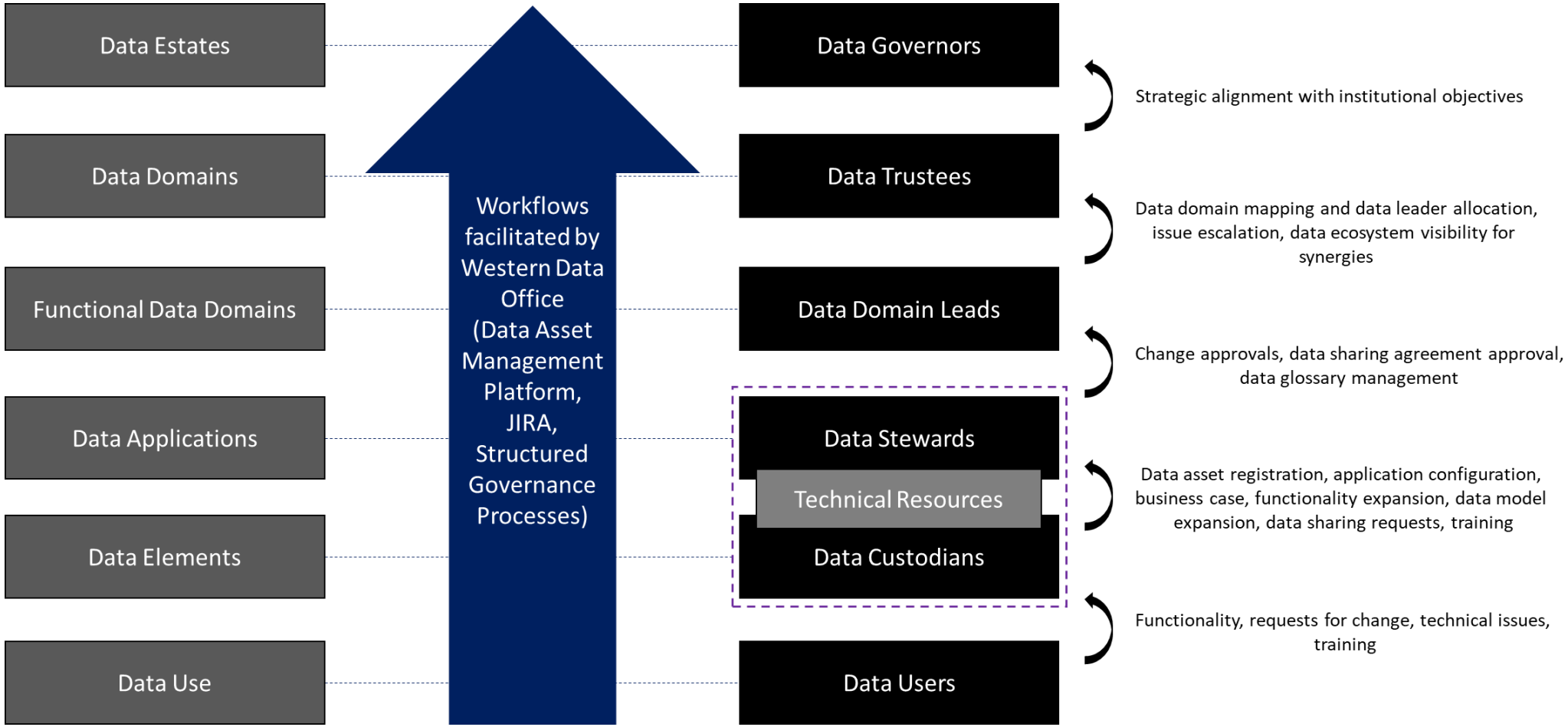
Data Requests and Data Sharing Agreements

Data requests occur multiple times a day across various contexts. It is proposed that a centralized data request process be established to ensure that each request is understood and filtered to the appropriate agents at the institution. Likewise, data sharing agreements would be formalized via the nature of the requests. Data requests can range from data catalog amendments to access to large institutional datasets to needs for data repository and analysis.

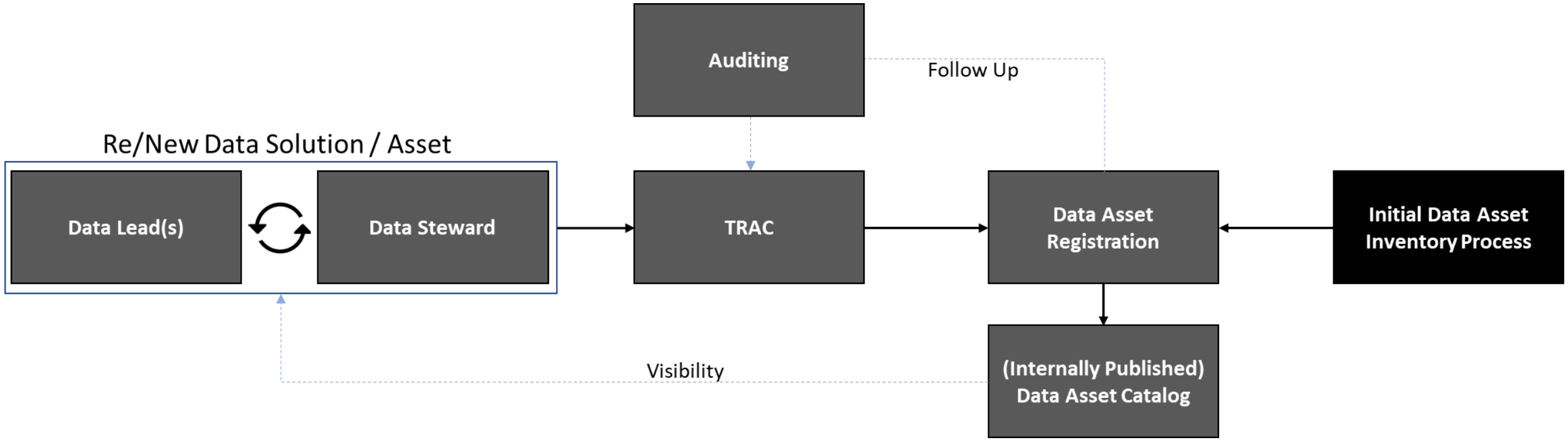
Deep Dives on Data Estates

To understand more deeply the various data implications at play and how interoperations function, a deep dive from data estate to data element will be performed. The deep dives will be a series of workshops designed to document the processes and definitions involved within a specific data estate. Relevant institutional partners will be assembled to assist in this activity.

How It Fits Together – Data Governance Workflow



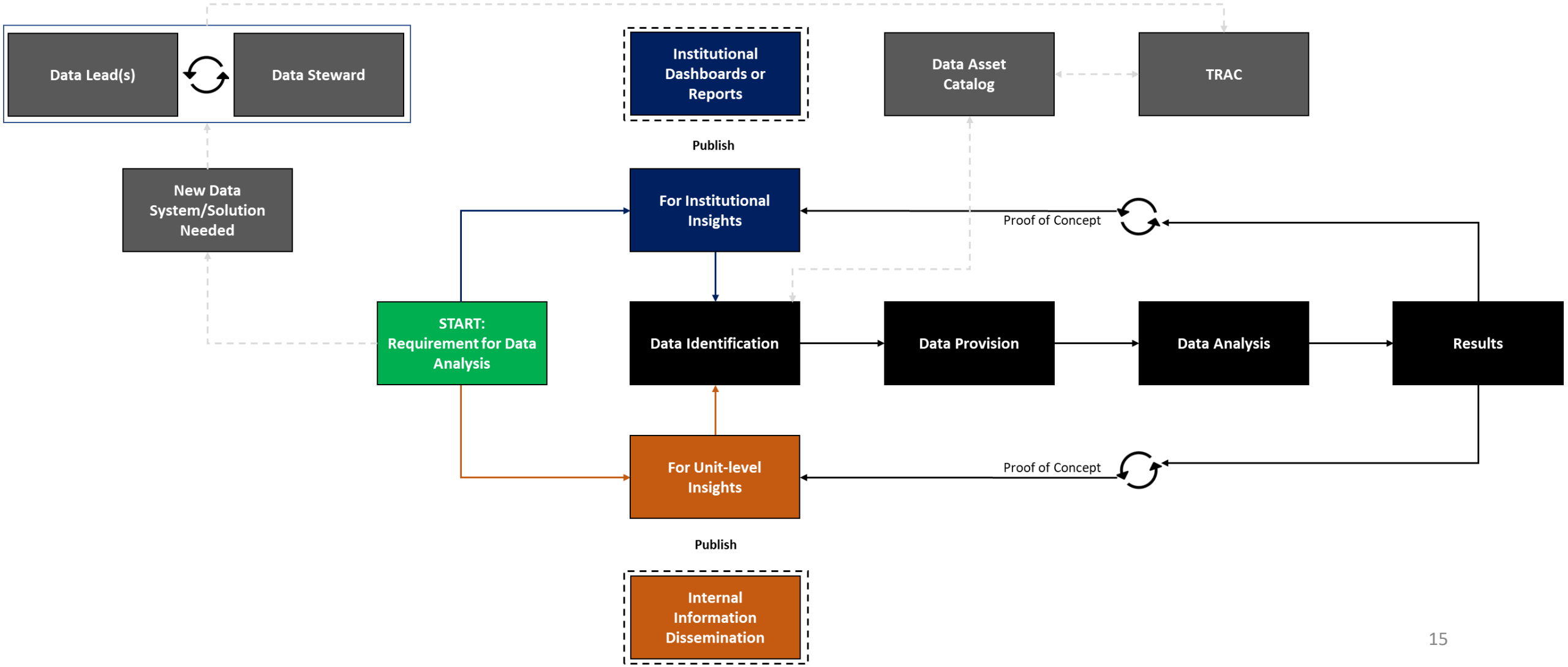
How It Fits Together – Data System/Solution Registration and Follow Up



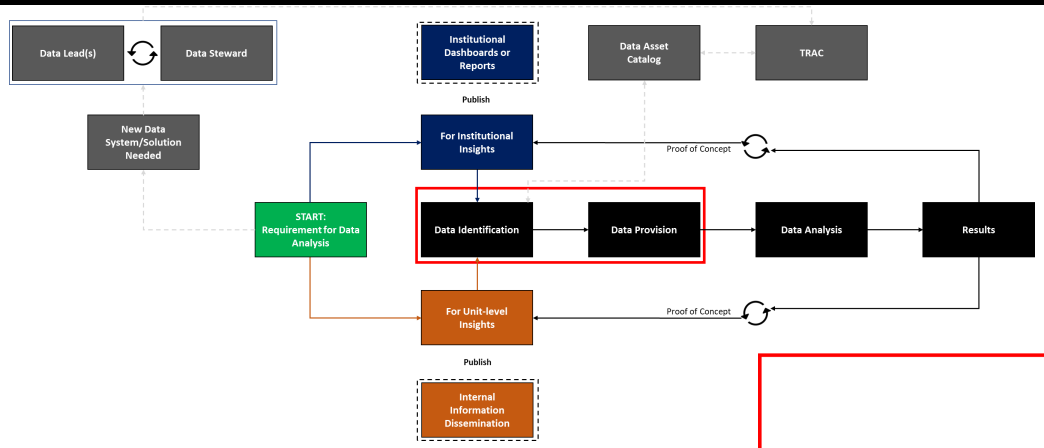
Data Asset Management Platform will allow Western to track Data Applications and attendant elements, including Data Stewards, Data Custodians, and Technical Resources.

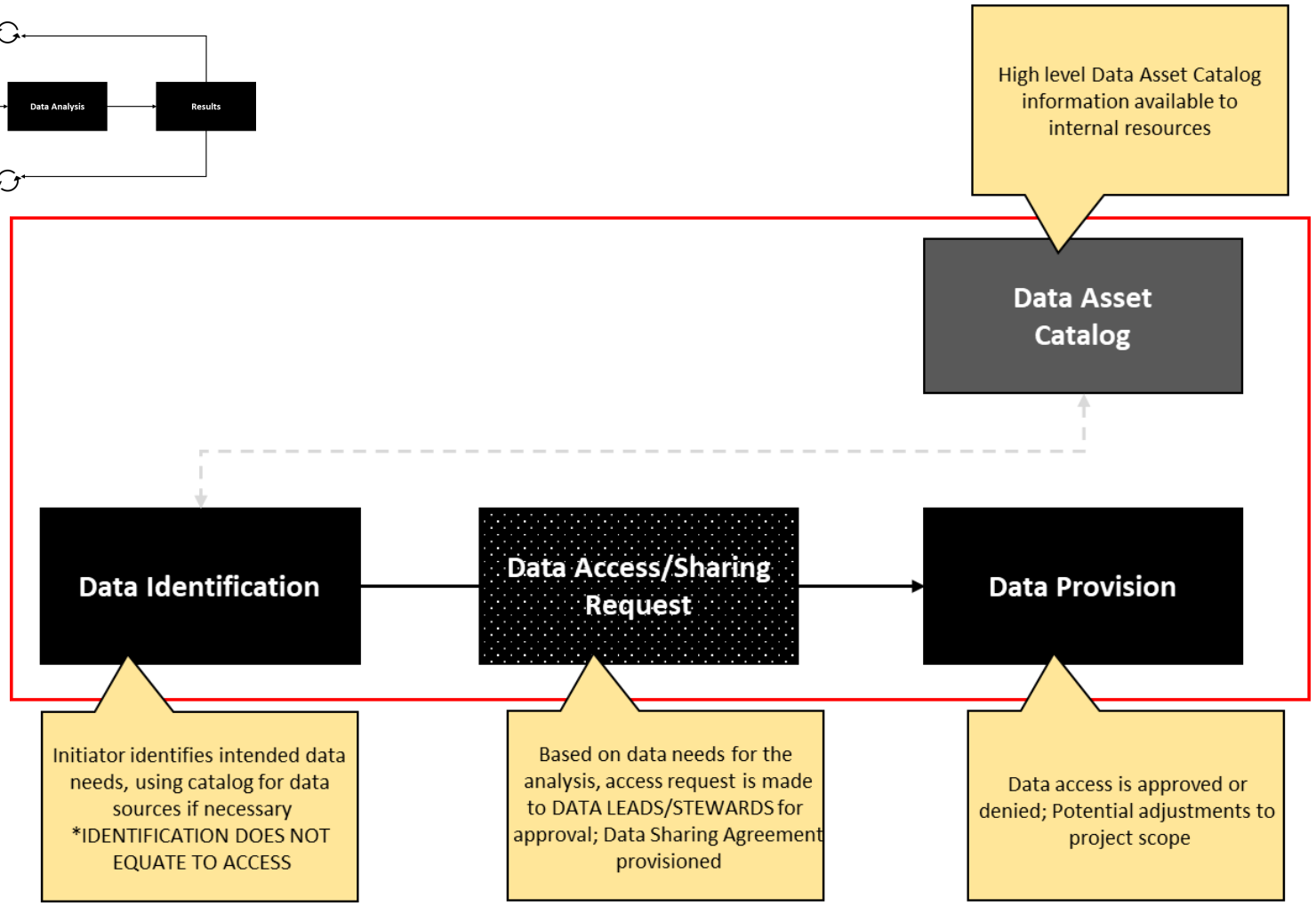
Data Asset Management Platform will also allow for mapping Data Leads with Functional Data Domains, providing a robust relationship map for approvals.

How It Fits Together – Data Source and Analysis Workflow

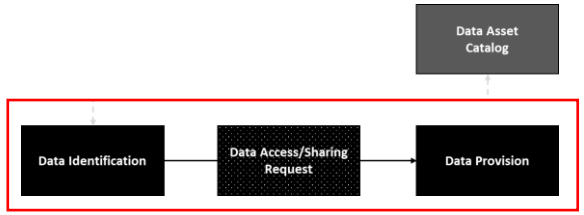


How It Fits Together – Data Sharing and Provisioning





How It Fits Together – Data Sharing and Provisioning



START: Requirement for Data Analysis

Data Identification

Data Access/Sharing Request

Data Provision

New Data System/Solution Needed

Institutional Data Asset/Source (*Systems of Record*)

Institutional or Unit-level Data Asset/Source I (*Line of Business*)

Unit-level Data Asset/Source II (*Local*)

Data Office

Data Sharing Agreement

JIRA Workflow

Data Leads / Stewards

Derived from FUNCTIONAL DATA DOMAIN and DATA APPLICATION mapping in Data Asset Catalog

How It Fits Together – Institutional Data Reporting