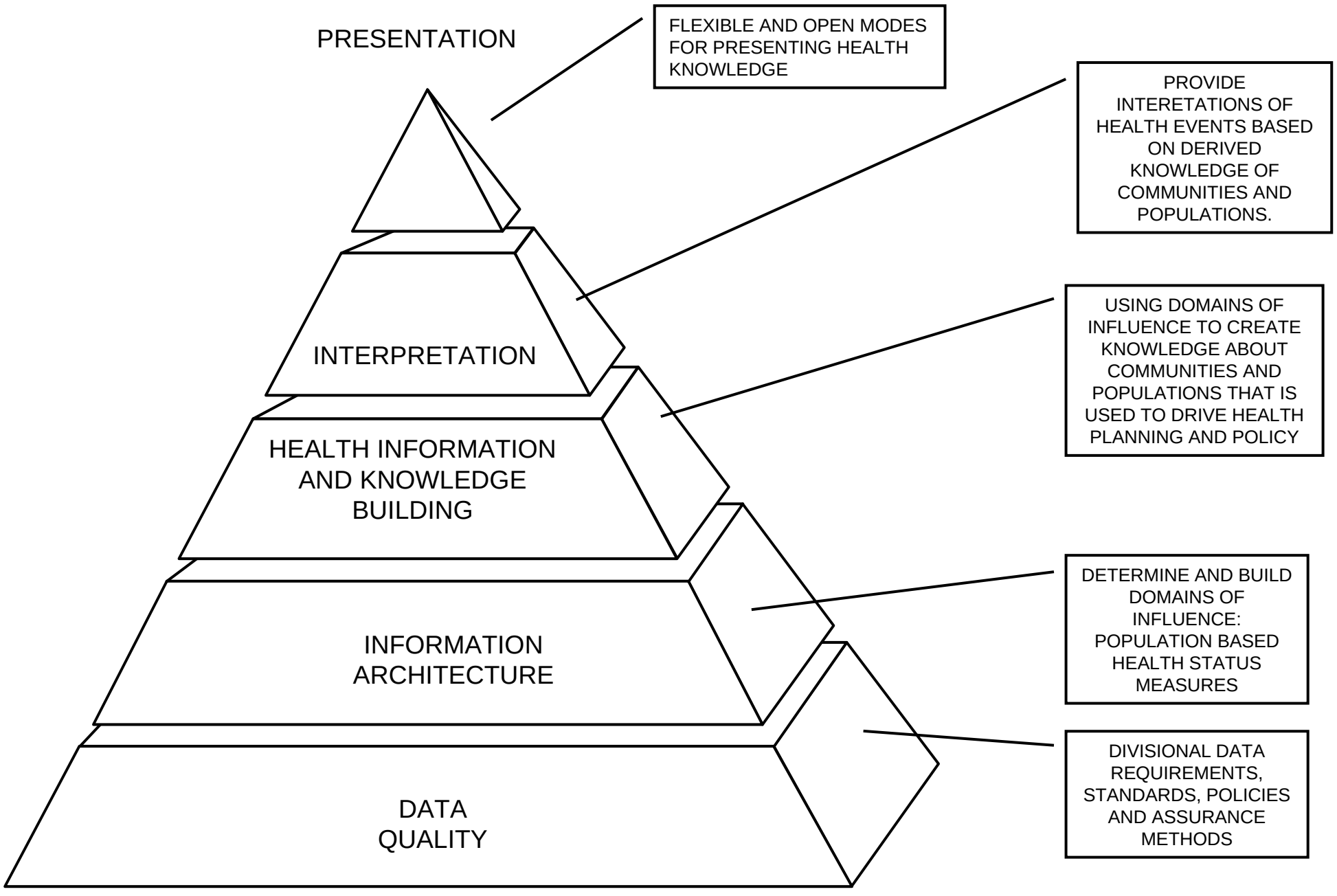
Foundation Goals of the Georgia Division of Public Health

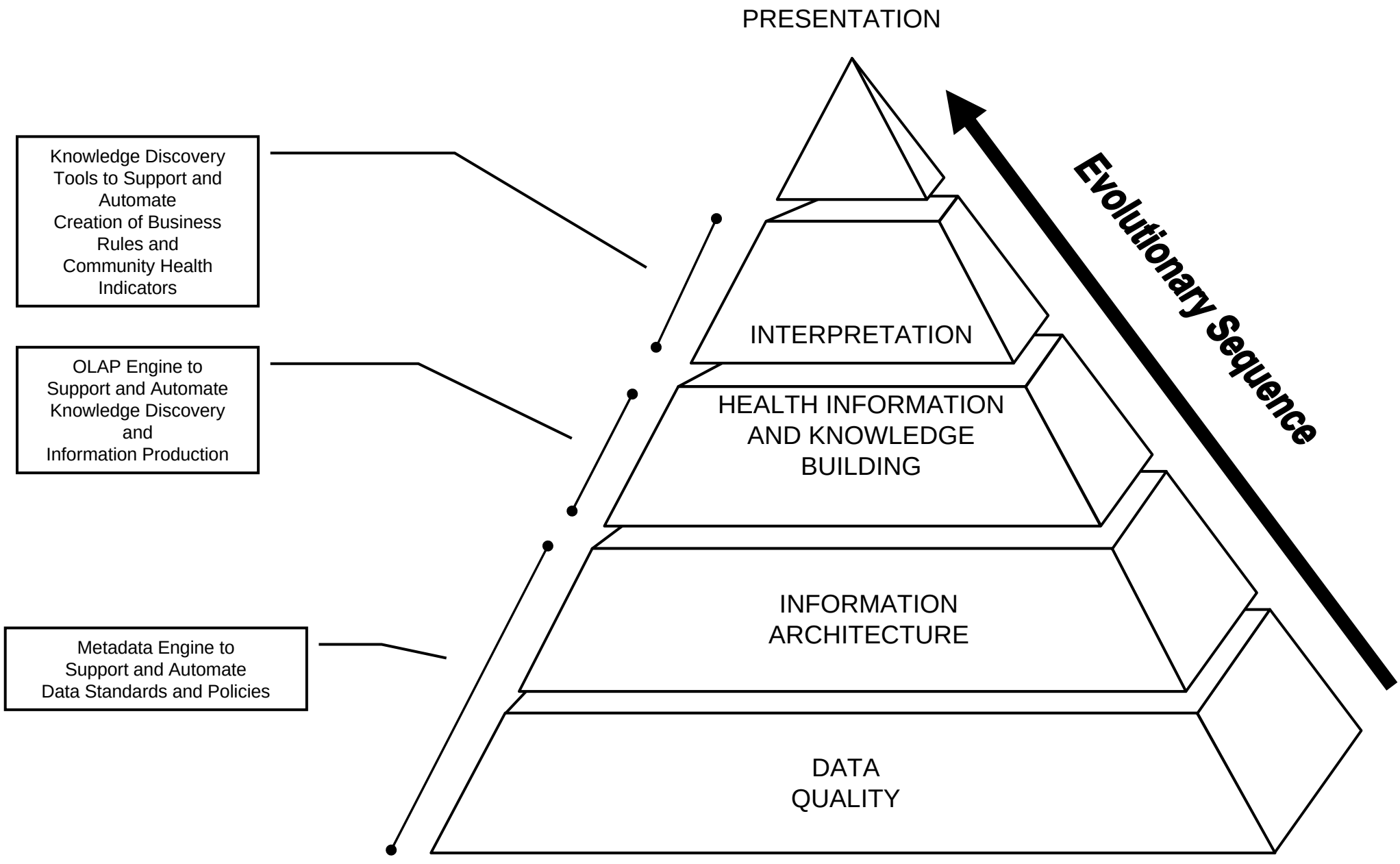
- **Assure valid and reliable health information.**
- **Provide relevant health assessment information products.**
- **Determine relevant community health indicators through analysis and data reduction.**
- **Create information products and knowledge bases that support and direct valid health programs and interventions.**
- **Provide training for users, staff and community leaders to use new forms of information products and health knowledge bases.**
- **Acquire staff that is competent in knowledge engineering, public health and information production.**

Assumptions

- **Public Health is a Process Control System.**
- **Public Health Management is Rational.**
- **Public Health Programs can have Specific Measurable Outcomes.**
- **The Rational Management of a Public Health System can be Accomplished by the Assessment of Prevention and Promotion Activities using Valid and Relevant Algorithms Applied to Data.**
- **Public Health is Data Driven -- Data are Collected that are Relevant to the Rational Management of a Public Health System.**
- **Data Collection Mechanisms are Agile, since Data Needs may Change.**
- **Public Health Systems are more Able to Act Effectively if Managed Rationally.**
- **Rational Public Health Management Reduces Organizational Complexity.**



Georgia Division of Public Health

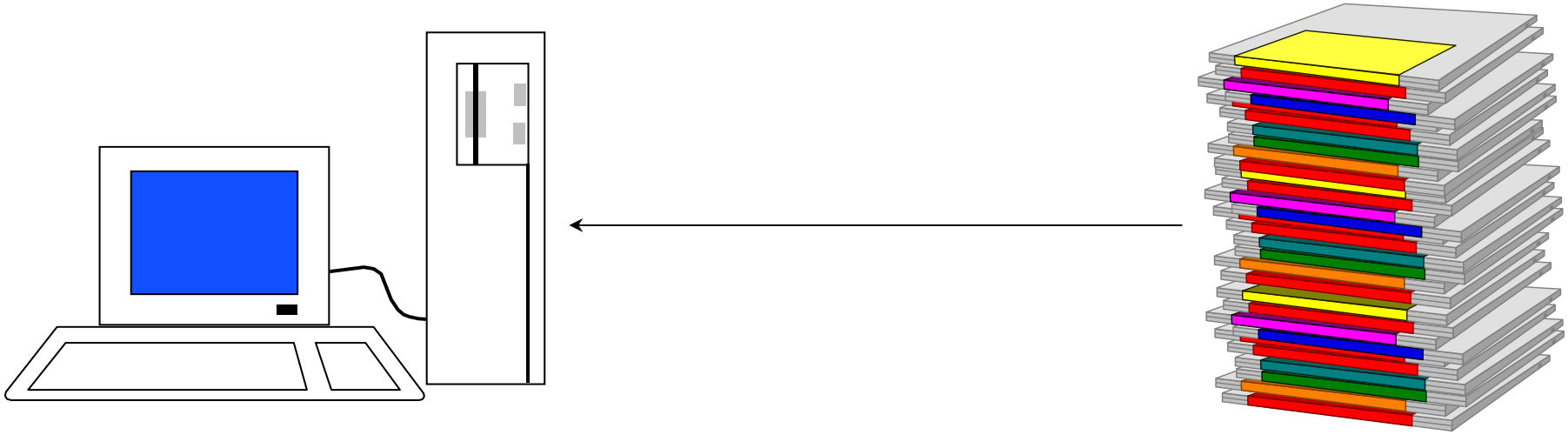


Georgia Division of Public Health

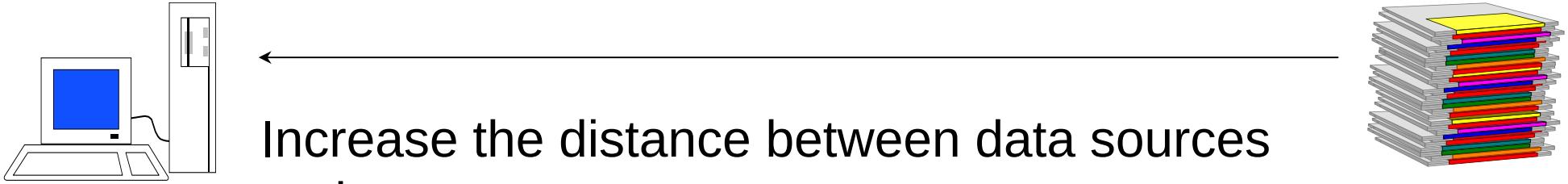
Definitions

- **Data Supplier -- an entity from which data are acquired.**
- **Incoming Metadata -- Methods and properties that apply data quality standards to supplied data, catalog supplied data into parts list, and load data into warehouse.**
- **Active Metadata -- Methods and properties for extracting and assembling cataloged data parts into information products.**
- **Information Product -- An assemblage of facts or knowledge that is relevant to the operation of the health system.**
- **Information Product Catalog -- A menu or list of available information products that are made available to users in a user interface.**
- **User interface -- A form of media (WWW, paper form, phone call, etc.) that allows a user to acquire information products.**

Currently users are close to data sources.



The GOAL is to



Increase the distance between data sources and users.

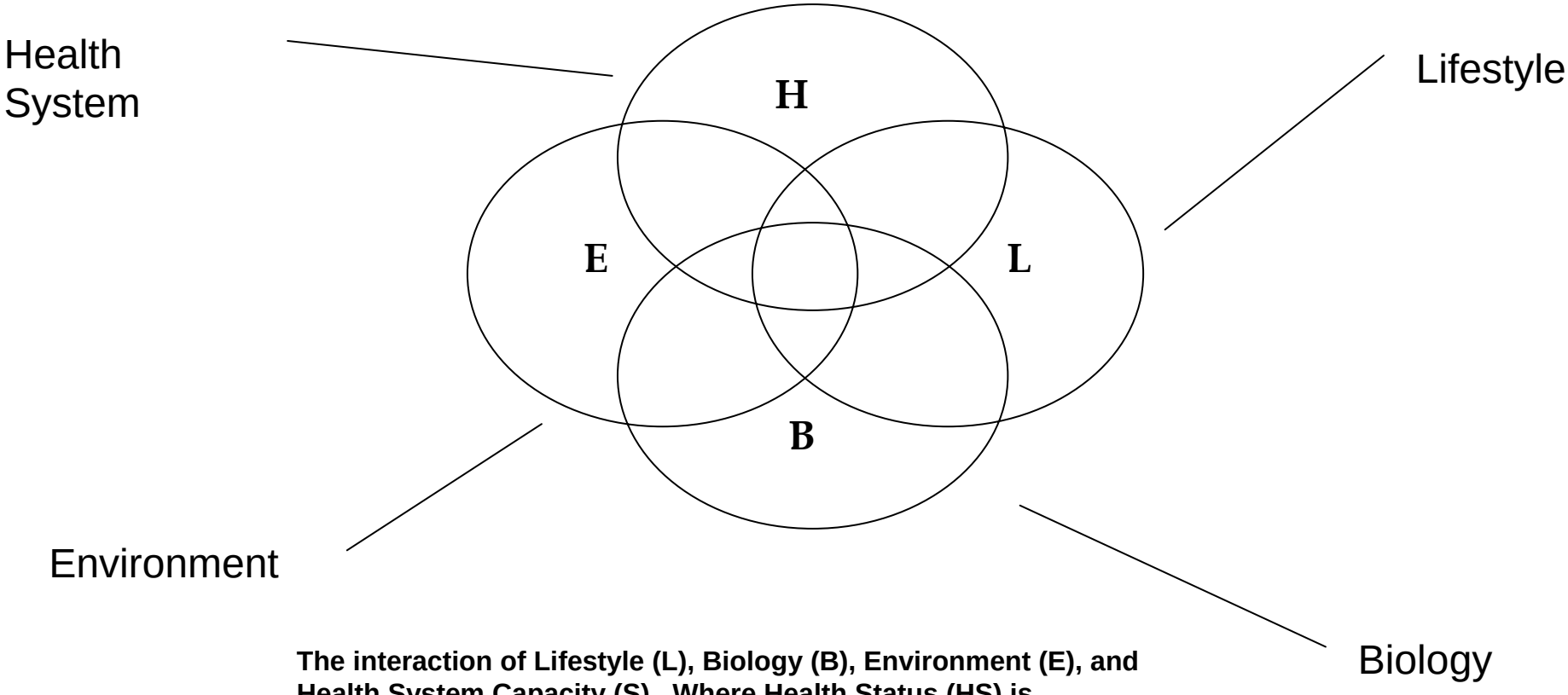
Distancing Users from Data

- Centralized data management.
- Assures consistent data quality.
- Assures reliability of information.
- Minimizes data fragmentation and redundancy.
- Users are provided relevant information products that are derived from data.
- Users focus on public health business functions instead of data management.

Progressive Model to Achieve Foundation Goals

- Build Health Information Data Quality and Architecture Based upon Health Field Theory
  - Complete in Year 1
- Build Health Information Factory upon Data Quality and Information Infrastructure
  - Identify Relevant Health Information Products that will be Produced by Health Information Factory
  - Assemble Health Information Products from Data Repository based on Data Quality and Information Architecture
  - Identify Methods and Modes of Product Distribution
  - Build Product Distribution System
  - Assure Consumer Satisfaction
  - Begin Year 1, Complete 1 half Year 2
- Build Health Knowledge Base to Drive Health Policy and Program Development
  - Determine Valid and Reliable Health Status Indicators
  - Distribute Health Knowledge Base through Health Information factory
  - Provide Policy and Program Analysis
  - Begin 2nd half Year 2, Complete 1st half Year 3
- Build Health Assessment Laboratory
  - Encourage External Investigators to use and Enhance Health Knowledge Base
  - Leverage Health Knowledge Base as a Training Technology
  - Provide Access and Training to Community Based Health System, Officials, and Activists
  - Begin 1st half Year 3, Complete 2nd half year 3

Health Field Theory

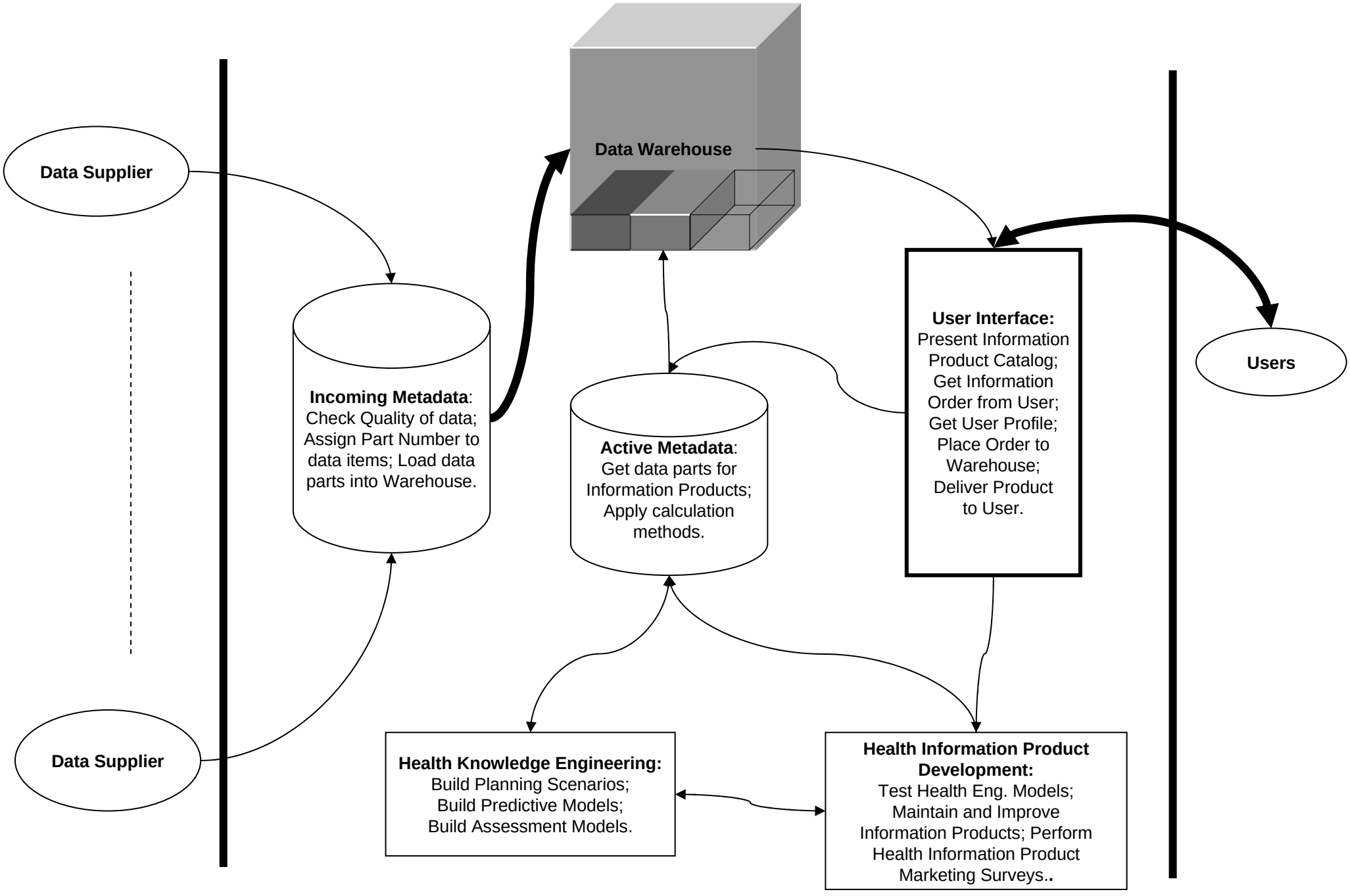


The interaction of Lifestyle (L), Biology (B), Environment (E), and Health System Capacity (S). Where Health Status (HS) is

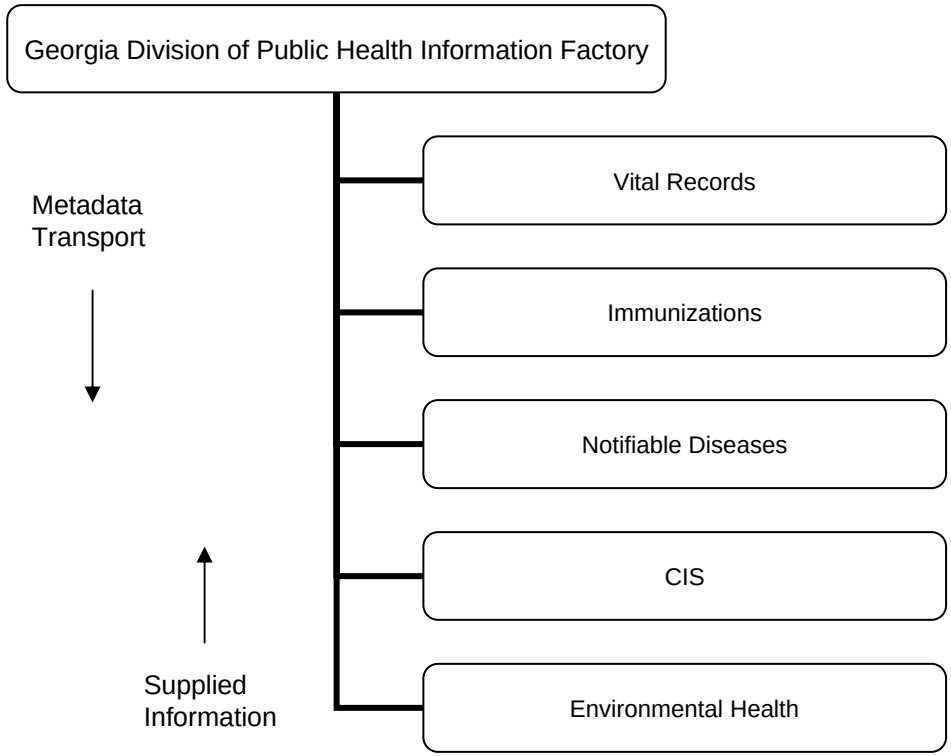
$$HS = E \circ B \circ L \circ H$$

Health Field Theory

- Represents the interaction of Biology, Environment, Lifestyle, and Health System.
- Assumes that this synergistic interaction both influences and expresses individual and population health status.
- Allows for multiple and multidimensional interaction configurations -- no one configuration may accurately represent or model all communities or programs.
- Allows health status to be viewed as influenced by components that can be used in simulations to predict health status, guide health policy and leverage public and private resources to improve quality of health support.
- Most important, health status need not be perceived as problematic or based upon disease, illness, or injury.
- Disease, illness and injury are independent variables of Environment.

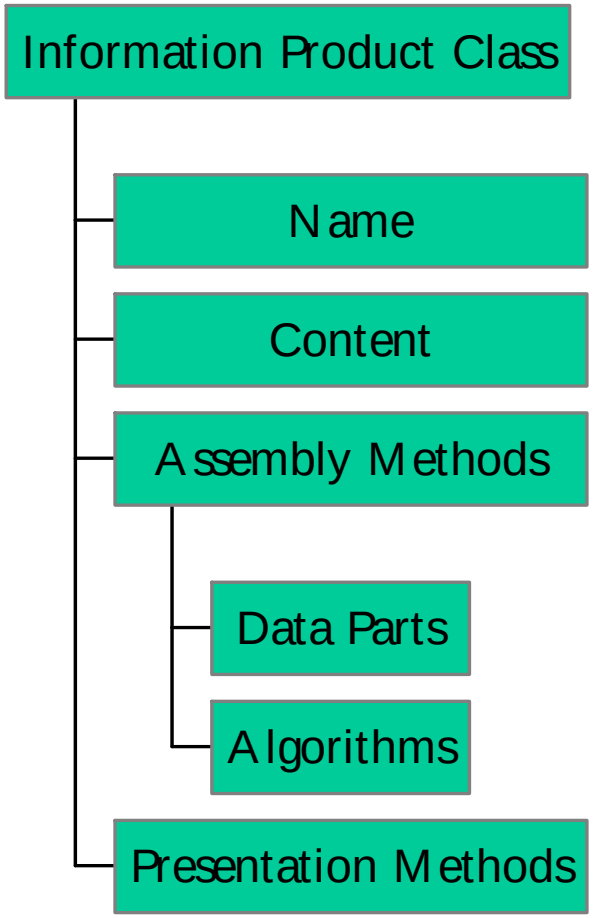


Georgia Division of Public Health

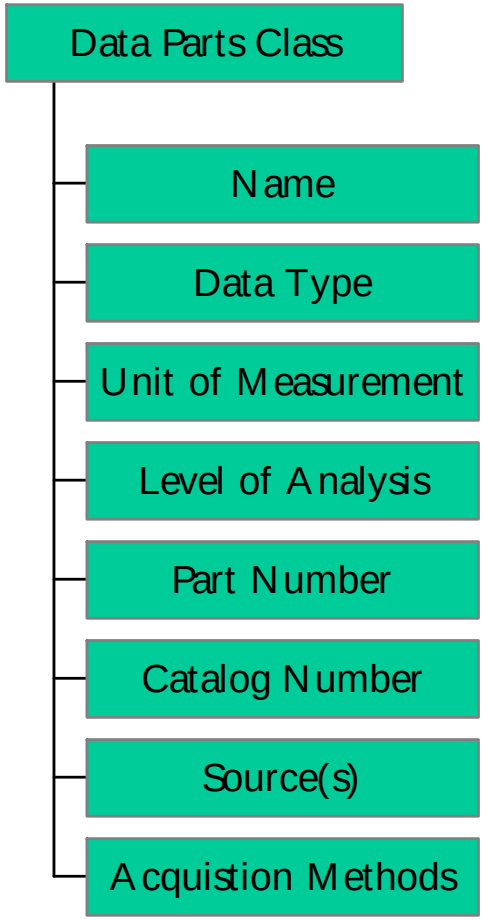


Georgia Division of Public Health

Information Product Object Model

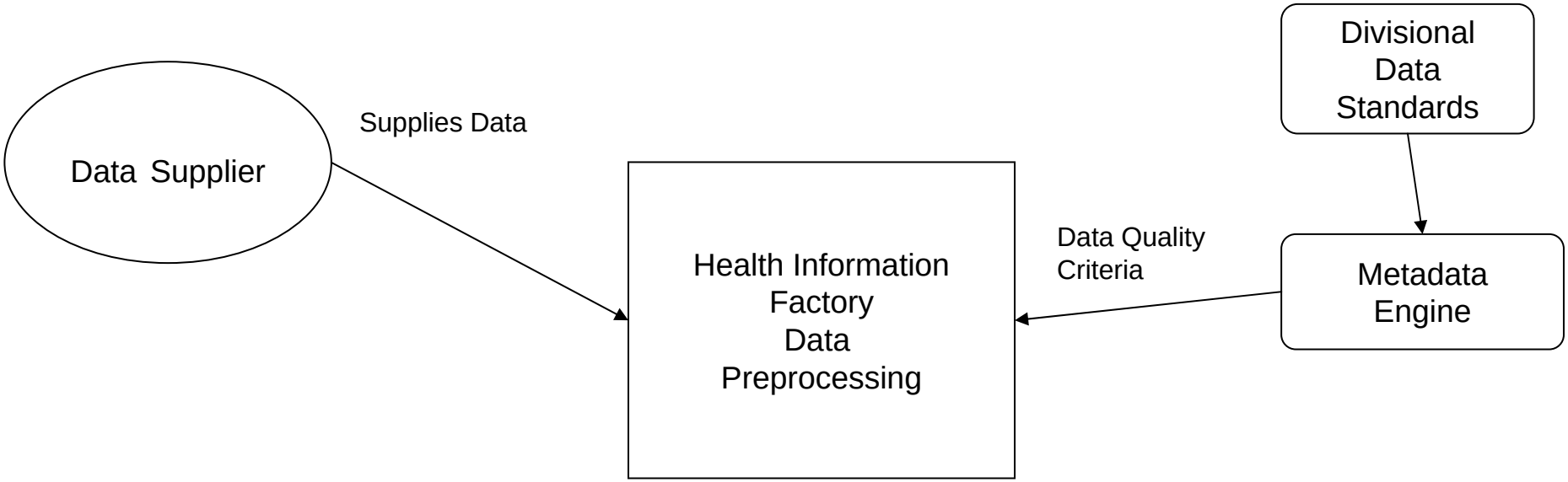


Data Parts Object Model: Data Specification



Health Information Factory Details

1 Data Suppliers and Data Preprocessing



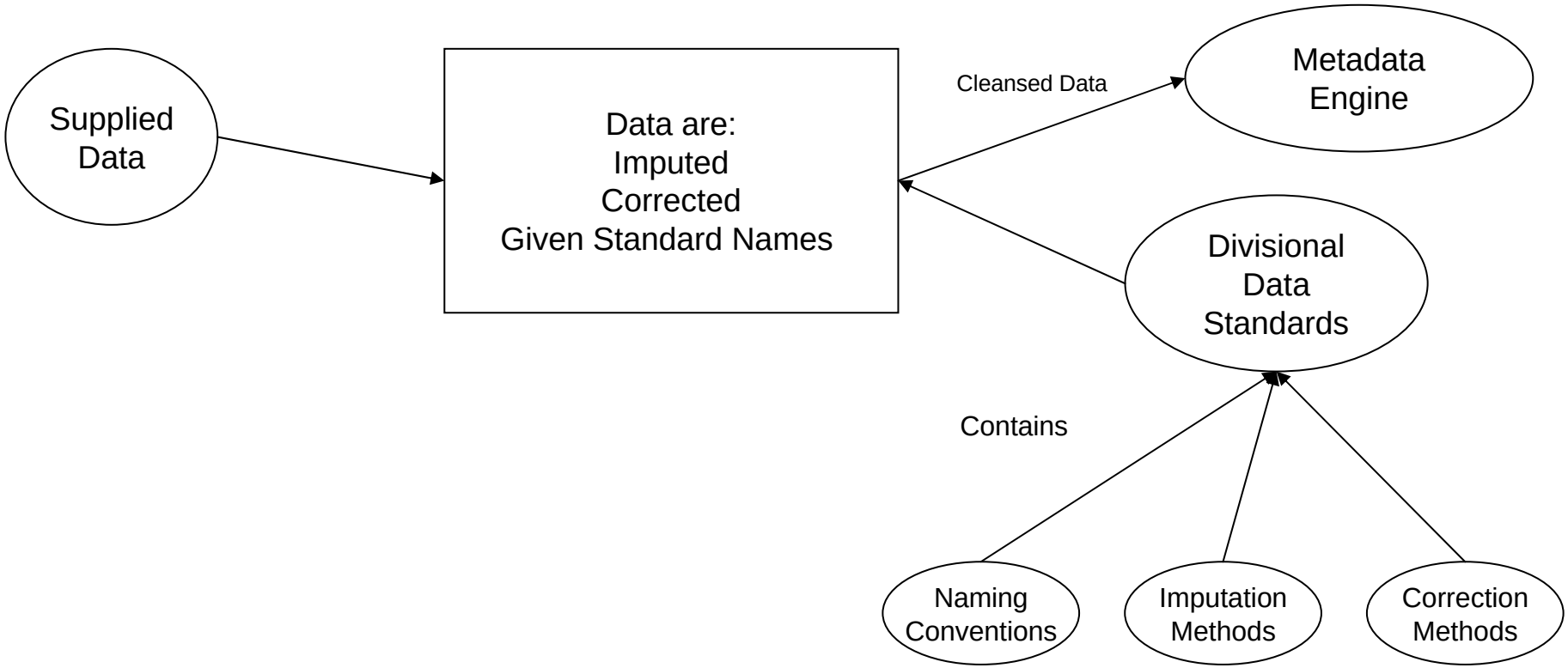
Health Information Factory Details

Data Preprocessing

- Data are acquired from data suppliers.
- Data are preprocessed using divisional data quality and consistency standards.
- A metadata engine contains:
  - data dictionary with variable specifications;
  - cross reference of variable usage;
  - methods for aggregating variables by multidimensions;
  - rules and methods for assigning variable specifications.
- Metadata engine enforces data quality and consistency standards.
- New data are updated in the metadata engine for distribution.
- New data are stored in a repository.

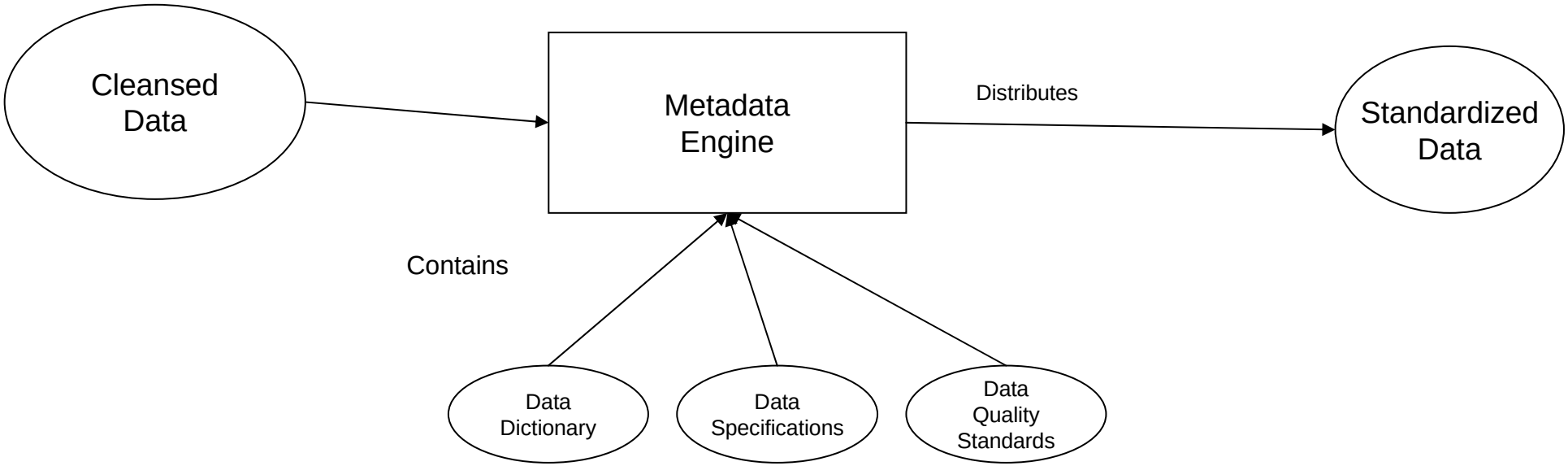
Health Information Factory Details

Data Preprocessing: Data Cleansing



Health Information Factory Details

Metadata Engine

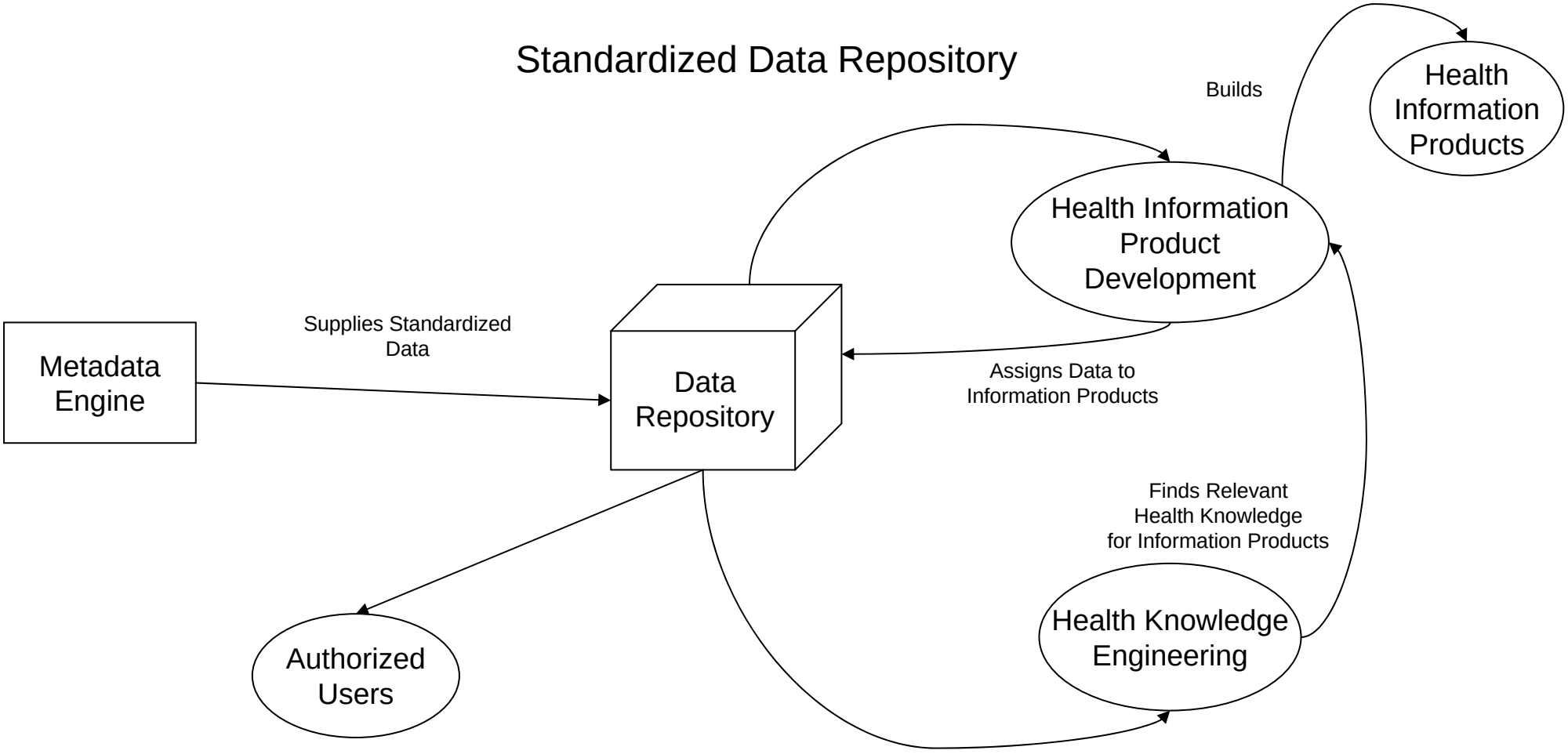


Health Information Factory Details

Data Cleansing

- Supplied data are cleansed:
  - missing values are imputed;
  - unknown and incorrect values are assigned consistent values;
  - out of range values are identified and labeled;
  - all values are manually checked, if necessary.
- Cleansed data are sent to metadata engine for further processing:
  - data types;
  - size;
  - unit of measurement;
  - presentation naming;
  - unit of analysis.

Health Information Factory Details



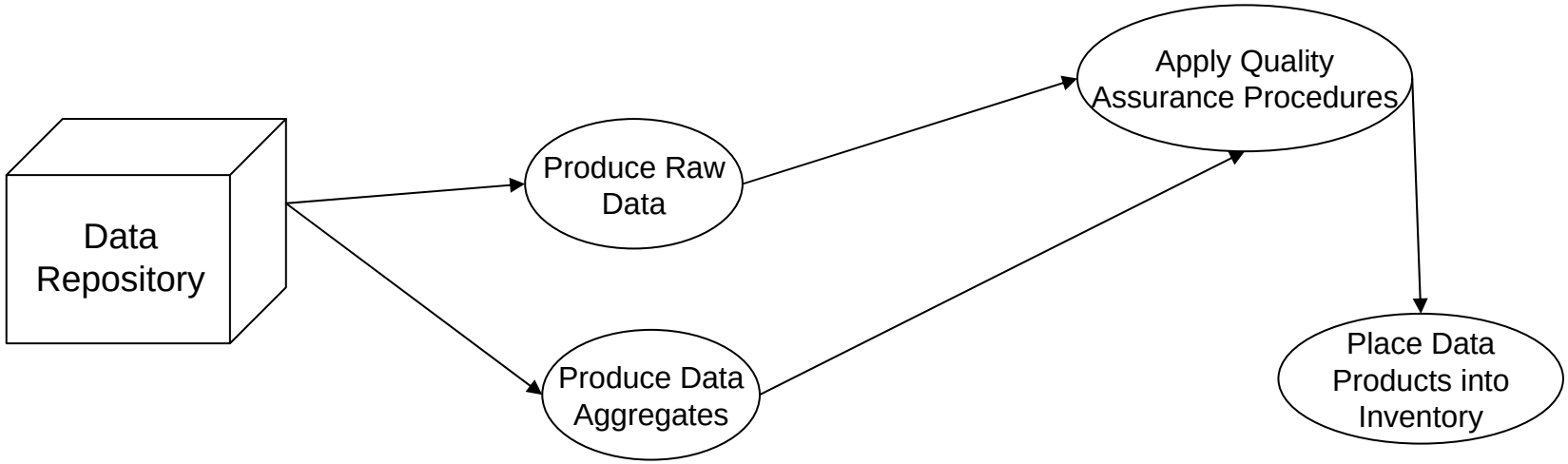
Health Information Factory Details

Standardized Data

- Standardized Data Repository contains:
  - Cleansed data;
  - Data items that meet divisional data specifications.
- Health Information Product Development provides:
  - Assignments of data items to information products;
  - Updates to Metadata, as needed, for new information products.
- Health Knowledge Engineering provides:
  - New health measurements and indices;
  - Updates of requirements to health information products.
- Authorized Users have access to standardized data for their own analyses.

Health Information Factory Details

Health Information Production 1



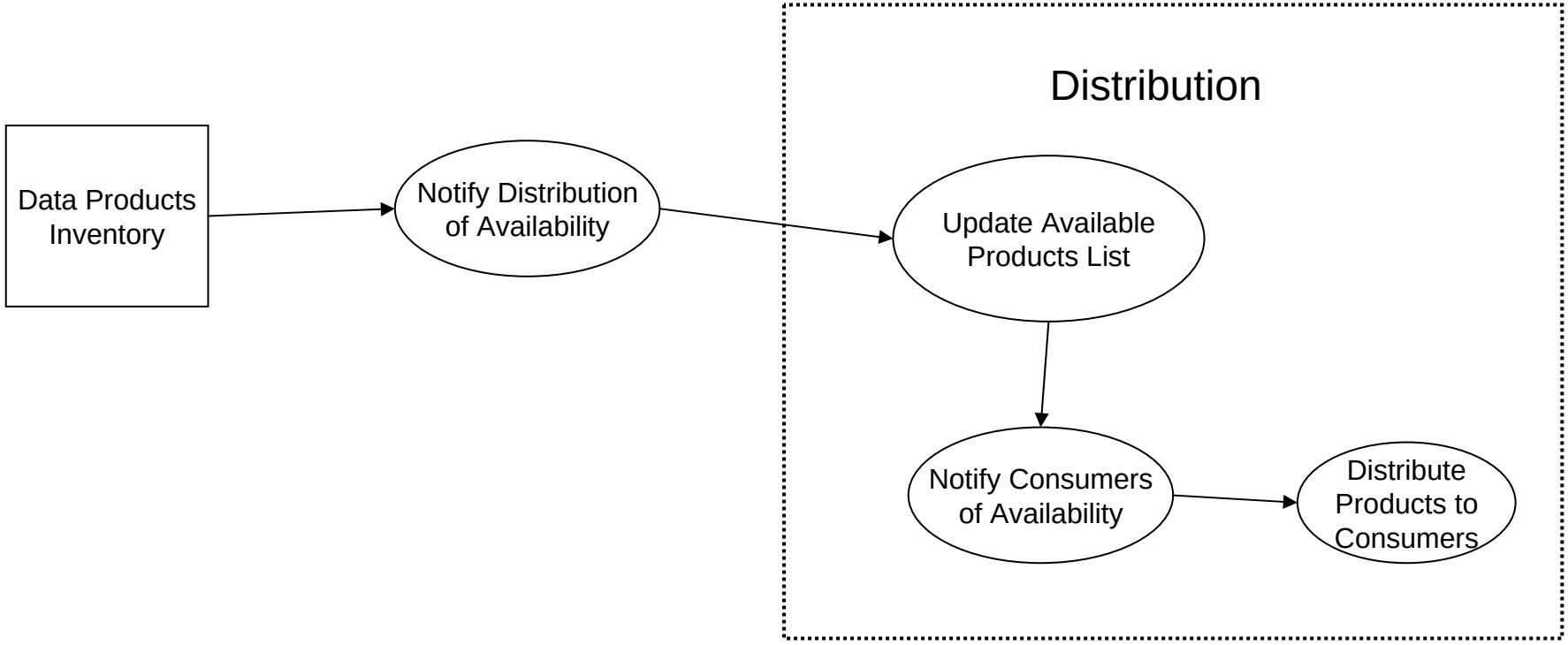
Health Information Factory Details

Health Information Production 1

- Get Data from repository.
- Produce Raw Data suitable for distribution to consumers.
- Produce Data aggregates suitable for analysis consumers.
- Apply quality control and assurance procedures to data products.
  - Modify data products, as necessary.
  - Certify quality of data products.
  - Certify methods of data production.
- Place data products into inventory.
  - Assign inventory number.
  - Notify distribution of availability.

Health Information Factory Details

Health Information Production 2



Health Information Factory Details

Health Information Production 2

- Notify distribution that new data products are available.
- Distribution updates available product list.
- Distribution notifies consumers of data product availability
- Distribution allows access to products by consumers.
- Distribution distributes data products to consumers in appropriate formats.