

# CORE DATA MANAGEMENT NEAR-TERM PRIORITY ACTIONS (2020-2024) DRAFT WHITE PAPER

## INTRODUCTION: FORMALIZING CORE DATA MANAGEMENT AT THE UW

Core Data Management (CDM), the UW's preferred term for what industry groups often call Master Data Management (MDM), refers to the curation of key entities that are used across the institution that provide a common point of reference. Core data is typically non-transactional data utilized across multiple systems and data domains. Effective core data management streamlines data sharing throughout the institution and can facilitate operational reporting and analysis across multiple systems, platforms and applications.

The role of Data Governance is to create a framework and define the policies and procedures in place to drive a successful CDM strategy. The ultimate goal of CDM as a whole is to provide the end user community with a "trusted single version of the truth" from which to base decisions.

## DEFINITION

As the UW's Data Governance groups draft a UW-specific definition of CDM, definitions from industry groups for Core Data Management provide a resource. According to the Data Management Body of Knowledge (DMBOK), Core Data Management (CDM) "entails control over Core Data values and identifiers that enable consistent use, across systems, of the most accurate and timely data about essential business entities".<sup>1</sup> The Gartner Group further describes CDM as the "workflow process in which business and IT work together to ensure the uniformity, accuracy, stewardship, and accountability of the enterprise's official, shared information assets."<sup>2</sup>

## APPROACH

The University of Washington (UW) will take an approach to CDM that focuses on processes rather than technical tools. This process-driven approach focuses on the organizational structure, people, and processes that are essential for creating CDM. Done correctly, overall benefits include:

- Reduced risk
- Operational efficiencies
- Simplified ETL processes and costs
- Improved data quality and accuracy

This focus directs the emphasis away from a single piece of technology or group and towards an emphasis on data themselves as the asset, more valuable than the various applications. This focus on data over technology or tools is important to the UW's approach to CDM, given that our data

---

<sup>1</sup> Knight, M., "What is Core Data Management (CDM)?", Dataversity (September 2017).

<sup>2</sup> White, Andrew and John Radcliffe. "Mastering Master Data Management." Gartner Research (May 2008).

ecosystem and budgetary considerations are not currently conducive to procurement of a central CDM tool.

## STRATEGIC PRIORITIES & GOVERNANCE FOR CDM

According to best practices summarized by industry consultants Gartner and InfoTech, the first step in firmly establishing or revamping CDM is to set the vision for the institution. The vision can help garner support from key stakeholders across the institution. Based on the vision, successful CDM strategies start small, demonstrate success with a limited scope and budget, and build out to scale. Attempting to overhaul systems and data structures will likely lead to limited gains.

Gartner and Infotech provide plans, processes, and templates that help organizations set their vision, strategy, and framework for CDM. Outlined below is a draft process for developing a vision, roadmap and guardrails for UW CDM based on common elements and recommended steps from these two groups with a greater reliance on the recommendations of InfoTech.

### LAY THE GROUNDWORK

Before work can start on developing robust CDM, there are several steps the UW will need to take (adapted from InfoTech).

1. Assess UW's readiness for CDM
  - a. Define data needs of the institution
  - b. Map current state
    - i. Get organizational buy-in
    - ii. Understand the existing data environment
  - c. Start with reference data, since this is a logical and often easier starting point
2. Determine business requirements
  - a. Create and adopt a UW definition of Core Master Data at the enterprise level
  - b. Take a targeted domain approach, determining which data domain contains the most critical core data in the organization so that, when ready, CDM work can focus on that domain

### SET THE VISION -- WHAT WE HOPE TO ACCOMPLISH WITH CDM

Data Governance groups can draft a vision for CDM that speaks to the following questions: 'What do we want to achieve?' and 'Why?' Based on the answers, the groups can develop goals, scope and high-level metrics.

1. Scope: What areas will be included in CDM?
2. Planning for the future of CDM: What is the proposed end-state maturity level of CDM?
3. Key metrics to measure success: How will we know we're meeting our CDM goals?
4. Key value: What positive impact will CDM result in for the UW?

## CREATING A CDM FRAMEWORK

KPMG provides a useful framework<sup>3</sup> for establishing strategic, tactical and operational goals for CDM work. The UW Data Governance Steering Committee has reviewed and adapted this model for UW purposes.

## UW'S CORE DATA MANAGEMENT FRAMEWORK WITH NEAR-TERM PRIORITIES



**Bold: Near-term priorities** | *Italic: Priorities for future consideration*  
Adaptation of [KPMG's Master Data Management framework](#) for UW purposes

The framework above serves as the deliverable for the first bullet point under Governance and the items in bold represent the UW's near-term priorities for the period from 2020 to 2024. The data governance groups recognized that the actions in italics, while not immediate priorities, are important and may become priorities later in a phased-out approach to CDM. This framework serves as the first step towards establishing UW's CDM. It provides both the framework and conceptual structure for CDM work, in lieu of a new tool or system. Below are details on the second step: providing an overall vision, strategy and guidance to core data management.

## PROVIDE OVERALL VISION, STRATEGY AND GUIDANCE FOR CDM

Infotech provides additional information on setting strategy. UW could tailor this work to our near-term needs and priorities.

1. Assess current CDM Capabilities
  - a. Measure the current state data governance surrounding CDM at UW
  - b. Develop future state goals for data governance around CDM at UW

<sup>3</sup> van der Staaij, A.J., "Master Data Management as a Global Business Service", Compact (a publication of KPMG IT Advisory) (January 2018)

- c. Determine which implementation style or styles would be appropriate for UW
2. Initiative Planning
  - a. Determine specific target states that are framed in quantifiable objectives that can be clearly communicated.
  - b. Perform a gap analysis of the current and future state strategy. This will be expressed in terms of gap closure strategies, and will help clarify the initiatives that will populate the road map.
  - c. Determine the relative business value of each initiative, as well as the relative complexities of successfully implementing them. These scores should be created with stakeholder input, and then plotted in an effort/transition quadrant map to determine where the quickest and most valuable wins lie.
3. Strategic Roadmap
  - a. Narrow down key areas within the strategy where an iterative approach provides the biggest benefit
  - b. Create a roadmap and plan for communicating to others

One of the first documents that could provide guidance for the CDM strategy is a CDM guardrails document. This could also provide some of the content for a roadmap and discussion of prioritization of operational and tactical deliverables. The draft in Appendix A could serve as the beginning of a guardrails document (adapted from the Core, Reference, and Metadata Guardrails developed for UW Workday Finance Transformation).

## TACTICAL PRIORITIES & PROCESSES FOR CDM

Once a strategic vision and roadmap are developed and vetted with stakeholders, tactical considerations include first establishing CDM roles and responsibilities and, second, focusing their near-term CDM work on essential processes.

### ESTABLISH CLEAR ROLES, TOOLS AND RESPONSIBILITIES

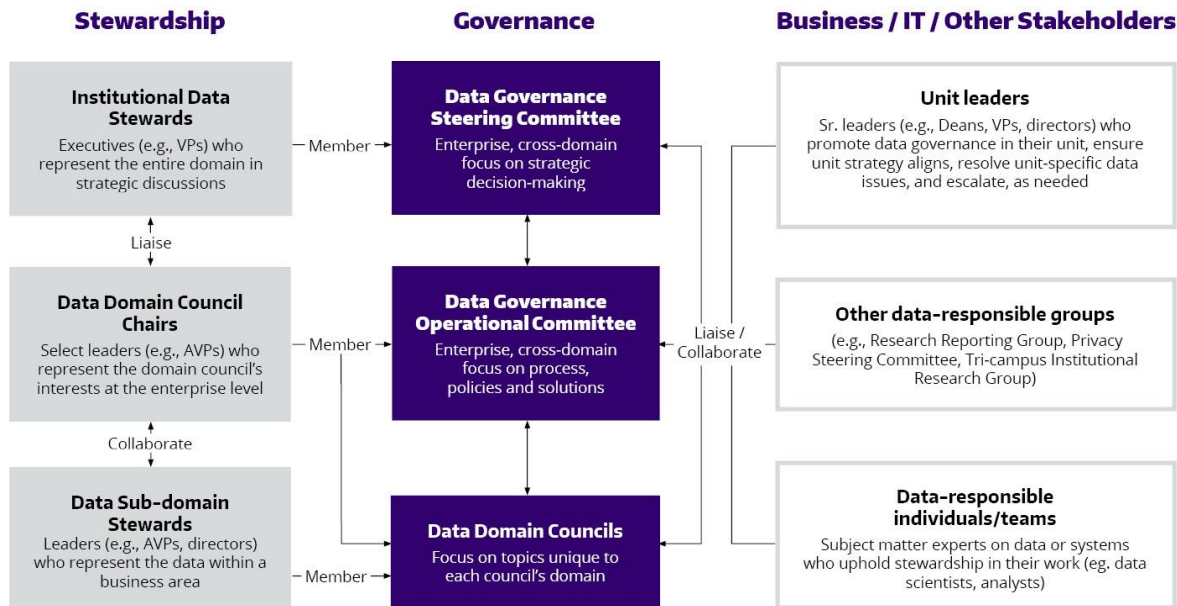
CDM is the center of a data ecosystem, and stewardship of CDM must be highly collaborative with other domain areas in order to be effective. The UW is currently revising and relaunching its data stewardship program. This work can be leveraged to articulate CDM-related roles and responsibilities and provide clear links between CDM, data governance groups as well as business, IT and other stakeholders.

The goals of the re-envisioned stewardship model are to:

1. Clarify that data are a University asset
2. Develop a viable stewardship model that
  - a. is user-friendly and easy to maintain
  - b. based on UW strategy and priorities
  - c. engages executives in strategic and tactical issues

3. Define data stewardship, roles and responsibilities
4. Define data domains
5. Set colleagues up for success through onboarding, training, and engagement
6. Create a consistent approach across and within core data and other data domains

The proposed UW stewardship roles and their relationship to other data-responsible groups are as follows:



The approach of data stewardship at the UW is to encourage data-responsible members of the community to think of data as part of an ecosystem in which we are dependent upon one another and rely on the thoughtful stewardship decisions and actions of others who have touched the data we need to do our jobs. The data stewards at various levels are part of and help anchor the people, process, and technology change. Their work, both overall and as it relates to CDM, incorporates the following elements:

- > Takes a collaborative rather than hierarchical approach, with stewards actively involved in creating domain-specific policies and processes based on frameworks, guidelines, and templates/tools provided by Data Governance.
- > Surfaces horizontal, enterprise-wide, cross-domain issues such as CDM as they engage in vertical, domain-specific discussions and decision-making.
- > Employs a RACI matrix to clarify roles and responsibilities.
- > Involves participatory governance and sustainable stewardship to avoid the common pitfall of hierarchical approaches that can break down over time.

The process by which the UW will gather input, vet and endorse this new stewardship model is as follows. Throughout the process, implications for CDM will be considered, including use cases related to specific core data. In Spring-Autumn 2020, the UW will continue to:

1. Develop and implement a data stewardship model
  - a. Construct a model with input from the data governance operational committee
  - b. Vet the model with key stakeholders and the data governance steering committee
  - c. Explore overlap with relevant APCs and address
  - d. Vet with the provost Implement, socialize and communicate.
2. Develop and implement a RACI matrix
3. Refine current policies on access to data (or clearly define who gets access to what data).

Once the revised stewardship model is launched and the domain councils are active, priority actions related to CDM will need to be established. Data governance groups may choose, for example, to prioritize the areas on the DMBOK wheel and determine the extent to which UW data governance should provide policies and operational guidance for essential processes, the next step outlined in the CDM framework. Or they may choose, for example, to focus on the finance data domain as Finance Transformation is rolled out, as an early example of CDM efforts in which there are opportunities to address CDM issues and needs. If so, this work could leverage the implementation of a new system as an opportunity to harmonize efforts across groups and to surface use cases that can be leveraged for the enterprise.

#### FOCUS ON ESSENTIAL PROCESSES (E.G. CURATION OF METADATA, DATA MIGRATION AND INTEGRATION, DATA MAINTENANCE, DATA QUALITY ASSURANCES & CONTROL, ARCHIVING)

Focused work on essential processes is more easily pursued once the stewardship model is implemented. After the stewards' roles and responsibilities vis a vis CDM are in place, focus can shift to developing and defining the essential processes necessary to support CDM at the UW. Recommendations for responsible groups to consider are outlined in the next section. These may be subject to change based on discussion and relative priorities and evolving realities.

## OPERATIONAL PRIORITIES FOR CONTENT & QUALITY, SYSTEMS & TOOLING

The work of operationalizing UW's near-term priorities for CDM rely on a main recommendation: Charge a group (or groups) to oversee all aspects of core data management at UW. While early recommendations included the creation of a Core Data Domain Council with representatives from each domain, in the near-term, it makes sense that the two data governance committees serve this role. Once membership is updated to reflect the proposed stewardship model, these groups will regularly convene senior representatives of each domain council. The chairs of domain councils will attend the Data Governance Operational Committee and each domain will be represented at the Steering Committee by an Institutional Data Steward. These two groups can work together to

prioritize CDM next steps and ensure progress towards them. This work includes establishing mechanisms within the larger data governance work to establish a CDM program structure including communications, meeting cadence and process to engage with the groups, as well as an intake process and a way to manage the backlog of work items. In the future, as CDM efforts mature, the data governance groups may wish to revisit the idea of a separate CDM-focused domain council.

Given the work to date by data governance task forces on core data issues, the committees (or a sub-set of members) may want to debrief with those groups dealing with core data so far (the Country Code task force and Funding Entity task force) and ask for their input and lessons learned. Namely, what support could they have used and how could their charge have been supported by a more formalized approach to CDM. Feedback from task force members may result in revisions to the proposed near-term next steps listed below for operationalizing CDM at the UW.

The following action items are listed in order of priority. They are meant as a draft to be updated and refined in consultation with data governance and existing task forces. They can be tackled directly by the data governance committees, in their CDM-responsible role, or by sub-groups or new task forces formed to address specific actions. Priorities listed below are grouped by topic and align with the operational near-term priorities of the UW's CDM Framework pyramid.

## CONTENT & QUALITY FOCUSED NEXT STEPS

### 1. Define categories of core data and conduct an initial inventory

Focusing first on defining categories of core data before other items will ensure that the space is well defined, and context understood. Defining categories of core data includes areas such as:

- Custom Naming
- Custom Mappings
- Custom Groupings
- Reference Data
- Core Data

(See Appendix B for detailed examples of each area)

The key deliverable is an inventory (as started in Appendix B) that includes the level at which each data element should be managed ("centrally," "regionally," or "locally"). A data inventory will ensure that locations and sources of core data are known, and overlapping governance structures are identified.

## 2. Establish quality standards for core data and its attributes

Develop a data quality assurance process that involves random sampling and procedures by which end users of data can report potential errors. Also included is the development of norms for each category, and processes for making exceptions and for adapting the norms.

(See [The Ultimate Guide to Modern Data Quality Management](#))

## 3. Enable standard core data definitions by formalizing the process and approval workflow

In parallel with the quality standards, the CDM-responsible group can work on formalizing the process and approval workflow for developing core data definitions. It is recommended that the group split this work into three separate action items:

- Formalize and codify the process by which core data specialist teams and/or task forces are created and charged, and how their work is formalized and approved (e.g., Country Code Task Force, Funding Entity Task Force).
- Second, define standards for delivery of core data by charged teams – what they are responsible for creating (e.g., terms, definitions, lists of valid values). Note that each charge might be unique in terms of what is expected to be created, based on the category of core data involved.
- Third, develop a process for reviewing, approving and implementing the outputs from these teams.

## 4. Provide guidance around data modeling and architecture (including metadata)

Provide guidance around data modeling, architecture and metadata, including formal recommendations, informal consultation, diagrams and data models. It will be important for the CDM-responsible group to clarify if this is a push or a pull model — whether it is a service that is provided (and therefore needs to be “advertised”) or whether it is a process initiated at the request of interested consumers.

## SYSTEMS & TOOLING FOCUSED NEXT STEPS

### 1. Track, inventory of where Core data lives, in which systems

Based on the initial inventory and definitions of core data mentioned above, the core data-responsible group can begin the work of developing more robust mechanisms for tracking core data across systems. Initially, this would result in a listing or spreadsheet of various “core data-like” entities, where that data is currently maintained, and current business and technical contacts for



that data. This work could be done simultaneously with other formative work given that progress is independent of the larger formative work and can be undertaken as time is available.

## FUTURE DIRECTIONS: LONGER-TERM PRIORITIES

After the near-term priorities outlined above are addressed, the data governance committees can take stock of the new, current state of CDM at the UW and initiate another round of prioritization and action. Some possible priorities surfaced in this work that were considered important but not near-term may be useful to revisit at that point. These include the italicized items seen below in the UW's CDM Framework.



**Bold: Near-term priorities** | *Italic: Priorities for future consideration*  
 Adaptation of [KPMG's Master Data Management framework](#) for UW purposes

In addition, the near-term tactical focus on essential processes will likely be ongoing, with the focus shifting to those essential processes not yet addressed. For example, developing guidelines for data migration and integration, and archiving processes.

Data governance may revisit the idea of a CDM-focused domain council and reassess the relative desirability/feasibility of CDM tools.

# APPENDIX A

## WORKING DRAFT CORE DATA MANAGEMENT GUARDRAILS

The following was adapted from the Core, Reference, and Metadata Guardrails developed for UW Workday Finance Transformation.

### Purpose

The University of Washington has a [Data Governance](#) structure that sets policies, standards and procedures for UW owned data. This document outlines the short and mid-term goals and strategies for core data management at the enterprise level, eventually resulting in policies and standards that will help align a distributed core data management model at the institution. This document provides guidance consistent with best practices and with assumptions about the future of reference and core at the UW.

### Scope

Currently, this document applies to the management of data within Enterprise Data Systems. For clarity, this document applies to all data sources and systems that are consumed by the UW's Enterprise Data Warehouse (EDW).

For purposes of this document, the **Enterprise Data Services** in scope are:

- **Enterprise Data Platform (EDP):** The enterprise data platform for UW Academy (UWA), which includes data stores and services for integrating with and reporting on Institutional data (including Enterprise Data Warehouse, Enterprise Integration Platform, and Enterprise Web Service).
- **UW Medicine Data Platform (UWMDP):** The enterprise data platform for UW Medicine (UWM), which includes data stores and services for integrating with and reporting on institutional data.
- **Identity and Access Management:** The systems that depend on Reference or Core Data in a System of Record to define groups or access and provide functionality to other systems based on this.

In future this list may be extended to include systems with similarly wide use and adhering to comparable standards for security, availability, operational support, business continuity, and ongoing strategic investment.

### Audience

This document is intended for use by:

- Data and Systems implementers

- Teams currently managing data in one or more of the UW's Enterprise Data Systems
- Teams currently managing core or reference data
- Teams operating enterprise data services related to institutional data
- Data Stewards
- Data Governance Steering and Operations Committee Members
- Unit and Central Data Analysts

## Visibility

Publicly readable and accessible to the world.

## Glossary

For data management terms that have not yet been defined by policy at the UW, this document references the DAMA International Data Management Body of Knowledge (DMBoK) for industry standard definitions (see [Reference Models](#) section).

## Data Management Concepts

- **Data Management:** “[T]he development, execution, and supervision of plans, policies, programs, and practices that deliver, control, protect, and enhance the value of data and information assets throughout their lifecycles.” (DMBoK).
  - The DMBoK divides data management into Knowledge Areas, shown in the [Reference Models](#) section.
- **Data Governance:** “[T]he exercise of authority and control (planning, monitoring, and enforcement) over the management of data assets” (DMBoK).
  - At the UW, Data Governance is a formal decision-making structure charged by the President and Provost (see the [Data Governance](#) web site).

## Reference & Core Data Concepts

- **Core Data:** “[D]ata about the business entities ... that provide context for business transactions and analysis. An entity is a real-world object (person, organization, place, or thing). Entities are represented by entity instances, in the form of data / records. Core Data should represent the authoritative, most accurate data available about key business entities.” (DMBoK).
  - Example: Data about courses at the UW (including course number, title, and description) used in a variety of business transactions such as registering students or generating transcripts.
- **Reference Data:** “[D]ata used to characterize or classify other data, or to relate data to information external to an organization” (DMBoK).
  - Example: A list of country codes, used across a variety of systems in different activities, such as identifying the location of a UW student, employee, donor, sponsor, supplier, or property.

- **Reference and Core Data Management:** “Managing shared data to meet organizational goals, reduce risks associated with data redundancy, ensure higher quality, and reduce the costs of data integration.” (DMBoK).
- **System of Record:** A system that is the authoritative data source for a given data element or piece of information. For purposed of this document, Systems of Record are the designated, authoritative source for Reference and Core Data.

## DOMAIN PRINCIPLES

### Reference & Core Data

#### 1. Manage Reference and Core Data on behalf of the UW

##### Rationale:

- Reference and Core Data are business assets, and require active management to be of value to the enterprise.
- Keeping Reference and Core Data **authoritative and accurate** enables processes and systems to standardize on them, thereby enabling data across the UW to be recorded consistently and used effectively.
  - To be **authoritative**, Reference and Core Data need definitions and stewardship, based on agreement between relevant stakeholders.
  - To be **accurate**, Reference and Core Data need management of changes and data quality, based on consistently following appropriate processes.
  - Effective Reference and Core Data require **agreement** among relevant **stakeholders**.
  - The Data Trustee(s) or Custodian(s) responsible for the relevant data domain(s).
  - Business and technical stakeholders working on design of the relevant System of Record and Enterprise Data Services (e.g. UWFT Process Transformation Teams and Workday configurators).
  - Business and technical stakeholders representing other major systems that are expected to use the Reference or Core data, or that contribute to the Reference or Core Data (e.g. business and system owners of research Sponsor data as Sponsor Core Data gets defined for Workday).
  - The Data Governance Operations Committee, to ensure the right UW stakeholders are included, ensure alignment, and ensure that decisions are systems-agnostic.

##### Guidelines:

1. Newly established Reference or Core Data entities should be **identified and defined** by agreement between relevant stakeholders.
  - a. Identify requirements for the Reference or Core Data.
  - b. Identify candidate Reference and Core Data for review by stakeholders.
  - c. Establish shared definitions of the new Reference or Core Data.

- d. Identify the specific data entities and attributes that will be managed as Reference or Core Data (typically a subset of the data in a System of Record).
  - e. Ensure that data entities and attributes have clearly defined distinct purposes; do not overload fields.
  - f. Ensure that the needs of downstream users (such as reporting users) are represented in defining the proposed Reference or Core Data.
2. Newly established Core Data should be harmonized across existing sources to the extent feasible.
    - a. Identify existing sources of data for the new Core Data.
    - b. Build agreement among stakeholders on how the new Core Data will be derived from existing data and/or will replace existing data.
  3. There should be clear business stewardship of each Reference and Core Data entity.
    - a. Identify one or more Data Custodian(s) responsible for each Reference or Core Data entity.
    - b. Identify any other stakeholders that need to propose, review, or be informed of changes to the data.
  4. An ongoing process should be established for making changes to Reference and Core Data.
    - a. Establish a transparent, repeatable change control process for requesting, reviewing, approving, and distributing changes to Reference and Core Data.
    - b. See the [Processes](#) section for general guidelines on defining this and other processes
  5. Ongoing processes should be established for ensuring data quality for Reference and Core Data.
    - a. Define the criteria for quality of the data.
    - b. Establish a process for validating data quality to these criteria on an ongoing basis.
    - c. Establish a process for submitting, reviewing, correcting, and communicating data quality issues.
    - d. See the [Processes](#) section for general guidelines on defining this and other processes
  6. As with any other UW data, Reference and Core Data must be managed in accordance with the UW's existing Policies, Standards and Guidelines.

## 2. Enable enterprise-wide use of Reference and Core Data

### Rationale:

- Systems of Record contain Reference and Core Data that is relied on by business processes and systems across the UW.
- At the time of this writing, the UW's de facto **architecture for Reference and Core Data** is to store the authoritative record in a System of Record. The UW does not operate a UW-wide repository for Reference and Core Data.
- At the time of this writing, the UW's preferred **integrations architecture** for common and re-usable data is the use of Enterprise Data Services (UWA Enterprise Data Platform and UWM Interoperability Platform). Therefore, this is the preferred approach for systems across the UW to obtain Reference and Core Data.

- See the [Workday Integrations Guardrails](#) for more guidance on integrations.

Guidelines:

1. Reference or Core Data should have an **authoritative record** in a System of Record.
  - a. Identify one System of Record as the authoritative record going forward for each Reference or Core Data entity.
  - b. If additional systems will be an ongoing source of data flowing into a Core Data entity, establish clear contracts for how new data from such systems is obtained, validated, and incorporated.
    - i. Example 1: If one or more other systems regularly contribute data to establish Core Data about new Awards in Workday, there should be a clear contract in place.
    - ii. Example 2: Academic HR systems outside of Workday establish Core Data definitions used in external reporting for academic employment data where Workday is the system of record; changes in that Core Data should flow on a scheduled basis into Workday.
2. Each team responsible for a System of Record should work with the Enterprise Data Services team(s) to publish Reference and Core Data for UW-wide use.
  - a. Define the requirements for systems needing access to Reference and Core Data.
  - b. Define the standard web services through which Reference and Core Data will be made available.
  - c. Define the standard reporting data models through which Reference and Core Data will be made available.
  - d. If necessary based on prioritized requirements, make Reference and Core Data available by other means such as file downloads.
3. Each team responsible for a System of Record should, in its documentation and change management efforts, refer people to the Enterprise Data Service(s) for access to Reference and Core Data.
4. Systems that obtain Reference or Core Data are expected to maintain the integrity of the data.
  - a. The authoritative Reference or Core Data should not be altered.
  - b. Reference or Core Data may be extended locally in a way that enables people to clearly distinguish the authoritative Reference or Core Data from additional local data.
5. When Reference or Core Data is identified for which there is no appropriate System of Record, a solution for recording and managing this data should be agreed upon.
  - a. Example: In the course of UWFT, UW Data Governance may elect to implement an Enterprise Organizational Structure; this data would support the goals of UWFT but would not be financial data.
  - b. The solution should support the guidelines for managing Reference and Core Data in this document, such as enabling the appropriate processes for change and data quality.

- c. The solution should be agreed upon among appropriate stakeholders, including the relevant business stakeholders and Enterprise Data Service(s) teams.

### 3. Enable impacted systems to transition to changed Reference and Core Data

#### Rationale:

- A major new System of Record -- or a major update to a System of Record -- usually entails significant **changes** to Reference and Core Data, and is dependent on corresponding **remediation** of data related systems, reports, and integrations managed by other teams.
- Teams responsible for impacted systems, reports, and integrations typically rely on the team responsible for the System of Record for the best available information about changes in Reference and Core Data.
- To reduce risk and long-term costs to the UW, it is important that impacted systems are able to remediate their data promptly and continue normal operations using the new Reference and Core Data.

#### Guidelines:

1. Each team responsible for a System of Record should incorporate in its project planning the **change management** efforts needed to help other teams find out about, understand, assess the impact of, and implement data model changes.
2. Each team responsible for a System of Record should develop **documentation** about data model changes as needed to support the work of other teams, such as:
  - a. To-be conceptual data model
  - b. Documentation of to-be entities, attributes, and definitions
  - c. Crosswalks or mappings between as-is and to-be data entities, attributes, and values
3. Each team responsible for a System of Record with changes that significantly impact Reference and Core Data for other systems should provide access to data sets that enable the transition from as-is to to-be data.
  - a. Mappings or translation tables should be provided to systems that need to remediate existing data in place.
4. In large changes with many impacted systems, it may become necessary to provide an ongoing translation service between as-is and to-be data models for systems that could not be remediated in time.
  - a. This type of service should be limited to the least possible data, for the least possible time, and to systems that are truly essential for ongoing operations and for which remediation truly is not feasible.
  - b. Before implementing this type of service, compare the full cost of creating and maintaining the service with the cost of remediating systems more quickly.
  - c. To ensure that the full ongoing cost is considered, a decision to go forward with this type of service will be made through UW Data Governance and/or UWFT Sponsors.
5. During ongoing operations and changes to Reference or Core data in a System of Record, the team responsible for the system should provide a way for other system owners to find out about changes (as described Principle 1).

# APPENDIX B

## AREAS OF CORE DATA, BRIEF DESCRIPTION

### Custom Naming

#### Organization Friendly Names:

- In the current FAS Finance system, “Organizations” are a core concept. However, due to mainframe system limitations, each organization is given a 22-character maximum ALL CAPS description.
- This description is not acceptable to display in systems and reporting, and without a single agreed-upon institutional “friendly name”, every office must come up with their own solution for meaningful organization descriptions.
- We are currently (April 2020) in the process of replacing the institutional financial systems with Workday. However, we should expect that Workday will not resolve this problem, and UW will still need a solution for displaying “friendly names” other than what appears in the new system.

### Custom Mappings

#### Mapping Organization to Student Degrees/Majors/Curriculums:

- UW currently has two organization hierarchies that run in parallel: The Financial Organization hierarchy and the Academic Organization hierarchy. Financial analysts think in terms of the Fin Org hierarchy. Chairs, advisers, and program coordinators think in terms of the Academic Organization hierarchy. A cross-walk mapping is needed to walk from financial to academic hierarchy in a managed and consistent way.
- The Office of Planning and Budgeting is responsible for maintaining a mapping for assigning Degrees/Majors/Curricula to the organization that “owns” the degree/major/curriculum using financial organization hierarchy (not academic). This mapping is updated every quarter.
- The Office of Planning and Budgeting is responsible for implementing UW’s Activity Based Budgeting which distributes money to departments based on the department’s instructional activity.
- ABB requires OPB to join student and financial data in a consistent and transparent way by mapping data from the Finance source system (Financial Organization) to data from the student source system (Degrees/Majors/Curriculums). This joining of data across subject areas, necessarily has to happen outside of the source systems.

#### Classification of Instructional Programs (CIP):

- CIP codes are assigned to majors and degrees to help the University/State classify which classes are of certain instructional type (STEM/non-STEM, for example). The list of CIP codes is published every 10 years, but UW (OPB) a) overwrites the CIP codes to reflect UW policies/needs, b) maintains the assignment of CIP codes to majors and degrees every academic quarter.



## Custom Groupings

### SAGE Sponsor List:

- SAGE is the institutional system for creating and submitting proposals for research grants. Proposals are submitted to sponsors for review. SAGE groups the hundreds of sponsors into categories that are meaningful for the SAGE transactional system (e.g. Private Industry, Foundation, etc.).
- The system and process owner, the UW Office of Research, is also responsible for federal compliance reporting related to grant money received by UW. The federal reporting requirements group Sponsors differently than the source system, and analysts in the Office of Research have a need to group sponsors in a way that aligns with federal reporting requirements.

### Expense and Revenue Groupers:

- The Financial Accounting (FA) Office is responsible for responding to the state with an annual compliance audit. The audit requirements group expense and revenue data differently than the financial source system at UW.
- Since the audit requirements don't align with how data is organized in the source system, the FA office has a need to create custom groupings of UW expenses and revenues in a way that aligns with the state audit requirements.

## Reference Data

### Student Data Management Office (SDMO):

- The SDMO maintains a set of reference tables that describe in human-readable format all of the core concepts of the student lifecycle that get represented in the SDB.
- Whereas SDB talks in terms of a student's enrollment status in code (e.g. 12 or 96), the SDMO reference tables describe these codes in human-readable form (e.g. registered or withdrawn).
- These human-readable descriptions are critical to the hundreds of academic report users who rely on them for meaningful reporting.

## Core Data

### Country Core List:

- The UW has a global impact in terms of research, teaching, and alumni. However, reporting on country's is an inherently fraught political exercise. *If* UW should display an entity's name (e.g., Taiwan) and *how* (e.g., Taiwan R.O.C.) is a decision that should be carefully made authoritatively by a data custodian, not ad-hoc by reporting analysts across campus.
- However, UW does not have a single authoritative list of country names to be used as a default for systems or reports. Instead, department-level custom lists proliferate.

Funding Entity (nee Sponsor) Core List:

- Sponsor management happens in Advancement, OR, Grant and Contract Accounting (GCA) and potentially other domains, but it is duplicative and out of sync. There's a need for a single Sponsor Core List that spans subject area and system.
- There is currently no solution to meet this need.

# APPENDIX C

## RESOURCES THAT HAVE INFORMED UW CDM DEVELOPMENT

### Industry Leaders

Dataversity: [Data Governance & Data Quality | News & Articles](#)

Gartner: [Creating an MDM Vision, Strategy and Road Map](#)

Infotech: [Develop a Master Data Management Strategy and Roadmap](#)

KPMG: [Master Data Management as a Global Business Service](#)

### Peer Institutions

Emory: [Institutional Data Management](#)

Michigan: [Institutional Data Resource Management Policy](#)

Minnesota: [Enterprise Data Management & Reporting](#)

NYU: [Administrative Data Management Policy](#)

Wisconsin: [Institutional Data Policy](#)