**RDaF Preliminary Framework Core**

(Appendix E in NIST publication **SP 1500-18**)

**Notes:**
1) In the Categories and Subcategories, “data” means “research data;”
2) Bolded words indicate input from the Stakeholder Scoping Workshop; and
3) A * at the end of a word or group of words indicates that a definition is provided in Appendix G of SP 1500-18)

<table>
<thead>
<tr>
<th>FUNCTION (Data Lifecycle* Stage)</th>
<th>CATEGORY</th>
<th>SUBCATEGORY</th>
</tr>
</thead>
</table>
| ENVISION                        | Data Governance* Structure | • Identification of Goals and Roles  
|                                 |          | • Data vision and/or **data policy**  
|                                 |          | • Data management value proposition  
|                                 |          | • Data management organization  
|                                 |          | • Value of data (quantitative or qualitative)  
|                                 |          | • Legal and regulatory **compliance**  
|                                 |          | • **Data quality** (including Trust and Certification)  
|                                 |          | • **Data privacy**  
|                                 |          | • **Data ethics**  
| Community Engagement            | Stakeholder community(ies) | • Communication with stakeholder community(ies)  
|                                 |          | • Interactions with other organizations  
|                                 |          | • **Cross-community engagement** (across domains and sectors)  
|                                 |          | • Inclusivity in interactions  
| Data Culture*                   | **FAIR data principles** | • Value of data  
|                                 |          | • Roles and responsibilities  
| Reward Structure                | For data management | • Value of **data workers**  
|                                 |          | • **Incentives and institutional credit for data sharing** and reuse  
|                                 |          | • Disincentives for data sharing  
<p>|                                 |          | • Human Resources (HR) involvement  |</p>
<table>
<thead>
<tr>
<th>FUNCTION (Data Lifecycle* Stage)</th>
<th>CATEGORY</th>
<th>SUBCATEGORY</th>
</tr>
</thead>
</table>
| **ENVISION (continued)**        | Workforce/Career Paths | • Workforce skills inventory  
• HR’s role in data workforce development  
• Data management training  
• Workforce preparedness in **new and advancing technologies**, e.g., HPC, AI, ML, and computation services  
• Promotional paths, continual training, and career development |
| Data Safety and Security         |          | • Safety and security assurance  
• Data inventory |
| Strategy                         |          | • Organizational data management |
| Data Risk Management*            |          | • Risk assessment  
• Risk mitigation and management |
| **PLAN**                        | Chain of Control | • Documentation  
• Communication within organization |
| **Economics and Costs of Planning** |          | • Decision-making tools for data, including cost-benefit analysis  
• Cost breakdown, i.e., calculation of costs by data lifecycle* stage |
| Funding Planning                 |          | • Models for **provisioning resources**, i.e., direct, overhead, or mixed |
| Data Objects                     |          | • Quantitative and qualitative data  
• Software, models  
• Instruments  
• Data publications*, journal publications  
• Presentations  
• Other |
| Hardware/Software Infrastructure |          | • Interoperability  
• Persistent instrument identifiers |
| Data Management Planning         |          | • **Data management plans** (DMPs)  
• Lifecycle considerations: living documents or static proposals? |
| Scientific Data Standards        |          | • Sources of standards  
• General, domain-specific |
| Assessment and Controls          |          | • Goals/definition of success  
• **Metrics** or metrics structure, tracking use and impact measures |
<table>
<thead>
<tr>
<th>FUNCTION (Data Lifecycle* Stage)</th>
<th>CATEGORY</th>
<th>SUBCATEGORY</th>
</tr>
</thead>
</table>
| **GENERATE/ACQUIRE**            | Sources of Raw Data* | • Generated In-house experimentally or computationally  
• Collected from external sources |
|                                 | Experimental Data Generation | • Specification and recording of instruments and associated metadata  
• Description and recording of measurement protocols  
• Methods for data and metadata capture and recording |
|                                 | Computational Data Generation | • Commercial and/or custom software  
• Methods for computational variables (metadata) capture and recording |
|                                 | FAIR Principles for Data Generated In-House | • Data born FAIR  
• Data made FAIR |
|                                 | External Sources of Data | • Data acquired FAIR  
• Identification, collection, and recording  
• Metadata harvesting |
|                                 | Community-Based Standards for Formats | • Standards development organizations/sources  
• General, domain-specific |
| **PROCESS/ANALYZE**             | Data Provenance | • Original authoritative copy  
• Version identification  
• Provenance of data derived from other data  
• Provenance of scientific records across all the individual outputs  
• Timestamping |
|                                 | Data Architecture | • Design  
• Security  
• Configuration management  
• Hosting and storage  
• Use of cloud |
|                                 | Software Tools | • Data lifecycle*  
• Management and analysis  
• Commercial and/or custom tools  
• System resilience and adaptability  
• Maintenance |
<table>
<thead>
<tr>
<th>FUNCTION (Data Lifecycle* Stage)</th>
<th>CATEGORY</th>
<th>SUBCATEGORY</th>
</tr>
</thead>
</table>
| **PROCESS/ANALYZE (continued)** | Scientific Workflow Processes and Systems | • Workflow tools  
• Laboratory notebooks, i.e., electronic, paper |
| Data Inventory | • Formats and standards  
• Catalogs  
• Interoperability (across instrument manufacturer file formats) |
| Data Modeling and Analytics | • Processes  
• Tools |
| Data Representation/Models/Structures | • Dynamic data  
• General, domain-specific |
| Data Curation | • Policies and processes  
• Manpower |
| Metadata | • Types of metadata  
• Responsible parties  
• Specification of metadata standards  
• **Linked data** structure  
• Persistent identification (DOI) |
| **SHARE/USE/REUSE** | Legal and Licenses | • Ownership of data  
• Constraints and encouragement for data use  
• **Intellectual property** rights/restrictions  
• Usage agreements/terms/licenses and required permissions  
• Terms of service  
• Data sharing agreements and licensing  
• Data citation* |
| Data Publishing* | • Repositories  
• Referencing data/digital objects from journal articles  
• Supplementary material  
• Data linking |
| Data Citation* | • Citation metrics  
• Citation impact |
| Internal and External Data Access | • Access internally, e.g., the data generator  
• Access externally  
• **Programmatic access**, aka Smart API  
• Data access vs. data visiting |
| Levels of Protection | • Unclassified but sensitive information, e.g., de-identification, enclaves  
• Security classification |
<table>
<thead>
<tr>
<th>FUNCTION (Data Lifecycle* Stage)</th>
<th>CATEGORY</th>
<th>SUBCATEGORY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of Protection (continued)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applications and Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Architectures for Application and Use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criteria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Sustainability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storage and Preservation of Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moving Data from One Service to Another across Organizations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retention and Disposition Schedules</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Preserve/Discard**

The end-of-use and end-of-life provisions for research data in an organization, including records management, archiving, and safe disposal.

- Protecting limited data/secure platforms/enclaves
- Data anonymization*

- Technologies for use and analytics, e.g., AI, ML

- Extensibility across communities, including machine-based interactions
- Capturing insights from ML and use of these to improve datasets for future AI applications
- Capturing data performance characteristics
- Location of data (e.g., relative to instruments, in the cloud, transient copies)

- Use and impact
- Data longevity and support
- Orphan datasets

- Media to store and preserve data
- Data back-up
- Data repositories

- Roles and responsibilities
- Moving data from one agency to another, e.g., from a funded research agency to an agency with a permanent repository
- Registration of repositories: roles and responsibilities
- Disciplinary archives

- Data archiving, i.e., what is kept and not kept
  - Decision processes
  - **End-of-life** issues
    - Example: Responsible party for keeping raw data* feeds
    - Example: Store (or not) raw data*, given the large amount of storage needed
    - Deaccessioning/End-of-life
    - Recognition of removed data (gravestone)