

# Data Management at the Grand Canyon Monitoring and Research Center

*by*

Michael J. Liszewski

Grand Canyon Monitoring and Research Center



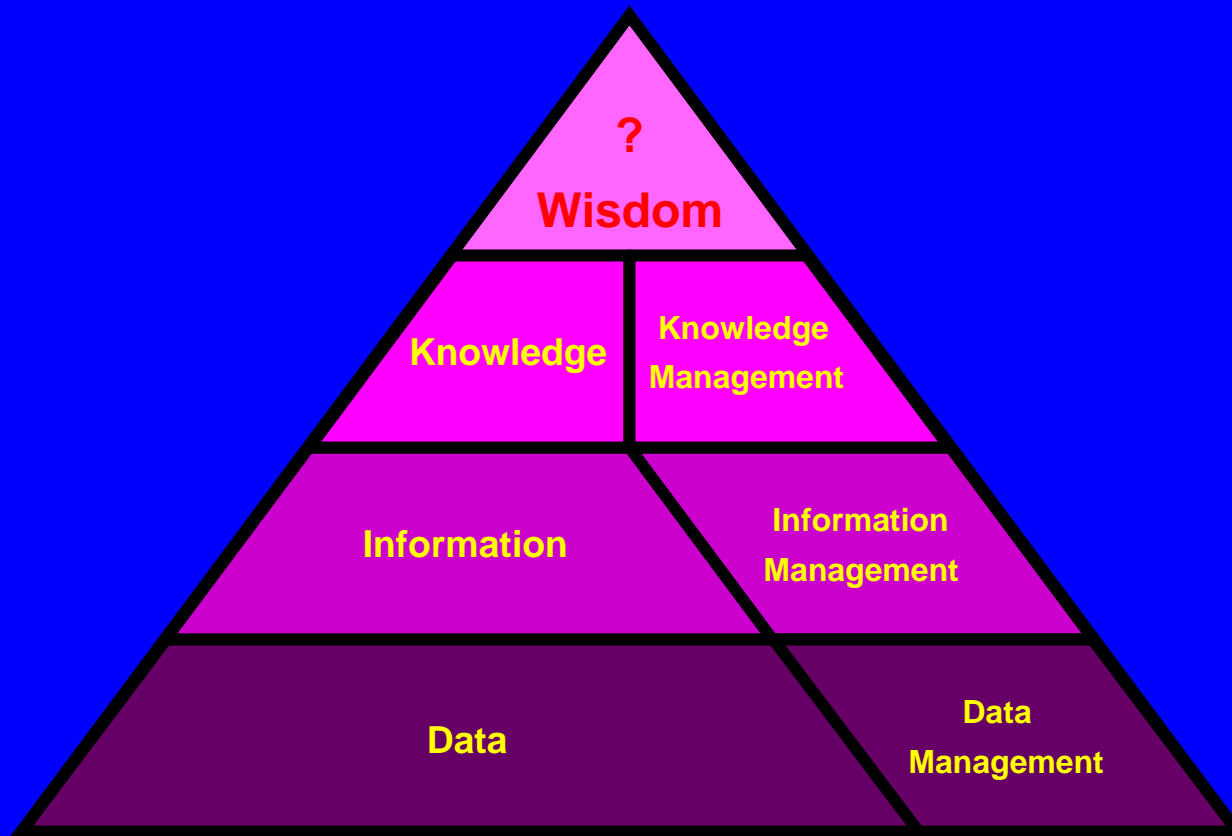
# Why data management?

1. Data is our largest and most worthwhile asset
2. Data collection represents our largest cost
3. Data is cumulative if collected and archived properly
4. Current data management practices will determine the degree to which we are able to use contemporary data in the future
5. Good management decisions depend on good data
6. Facilitates Interoperability
7. Stimulates new thinking and fresh ideas
8. Helps manage complexity

# Role of data in an adaptive management program

1. Provides information for management decisions
2. Facilitate interoperability

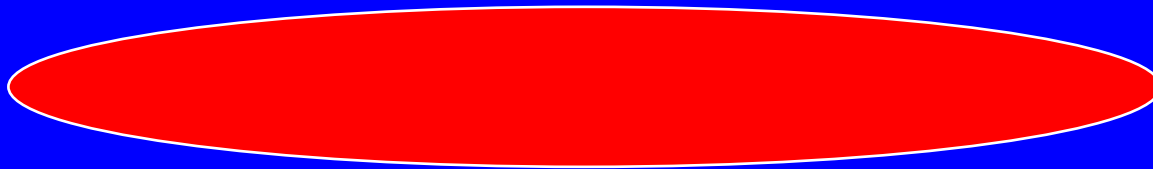
# Wisdom Triangle



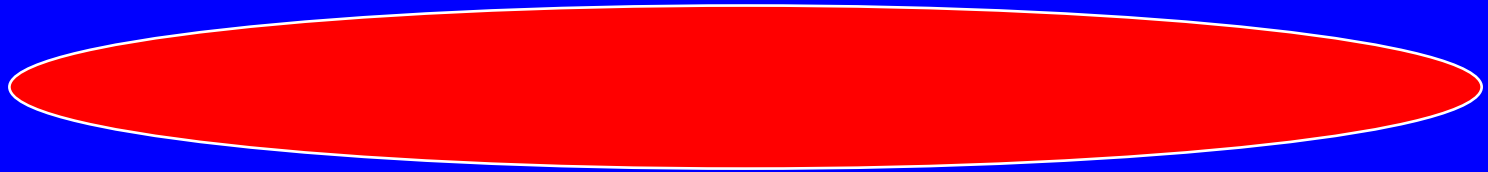
# INFORMATION INTEROPERABILITY

Defined as ....

**INFORMATION MANAGEMENT**

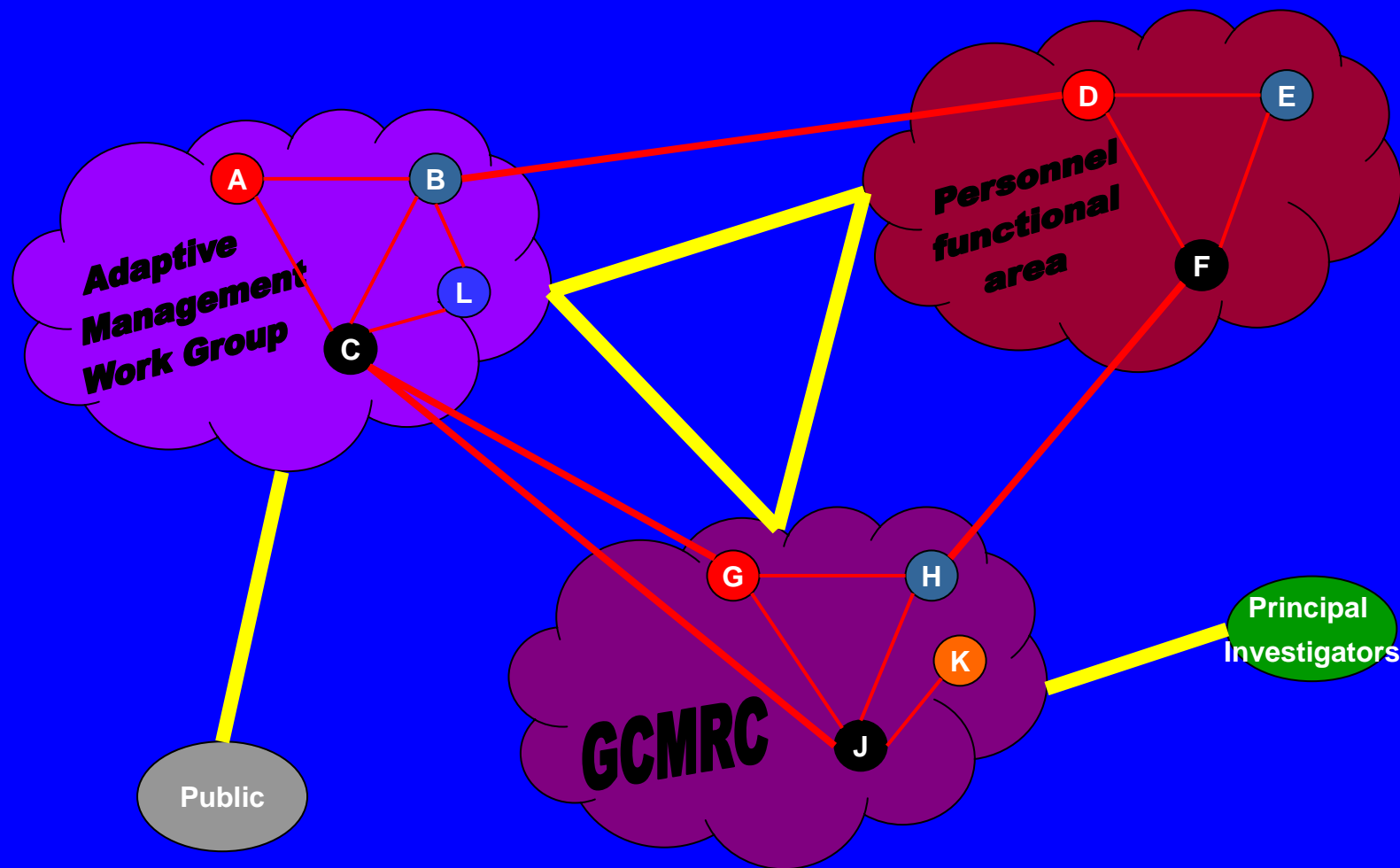


**combined with**



**DATA MANAGEMENT**

# Data clouds



# Fundamental data management requirements

1. **Integrity** (accurate, consistent, minimal disparity)
2. **Commonly understood meaning** across the organization
3. **Accessible** (searchable and retrievable)
4. **Longevity**

# Information management components

Requirement	Solution
Integrity	Data standards
Understood	Data definitions
Accessible	Infrastructure (DBMS)
Longevity	Documentation



# Data integrity - accuracy and consistency

1. Is it measurable at the level of interest?
2. Is it quantifiable?
3. Is it repeatable?
4. Does zero equal zero?
5. When and where collected?

# Data integrity - data disparity

Sample collection method

Sample analysis method

Accuracy and precision

Units of measurement

Classification of results

Space and time

Storage media, file format, and compression

Geoids, datum, and projections

Computer hardware and software

Etc., etc., etc.


# Data disparity must be removed

At the source – conformance to data standards  
After the fact – GCMRC staff time and effort


# The importance of data standards

1. Integrity
2. Consistency
3. Minimize disparity

# Data definitions



**ACCORD**  
**Data Definition Repository**



[EXIT](#)

[Home](#)

[DDR](#)

[Working](#)

[Cycle](#)

[Returned](#)

[Reports](#)

[Print](#)

[Help](#)

**Search**  
[Quick Search](#)  
[Advanced Search](#)

**DATA DEFINITION**  
Approved  
Label: **DATA DEFINITION**  
Description: A DESCRIPTION which determines the rules to which one or more collections of data instances must conform.  
Originating DMG: Central Data Management Authority  
Reference Source: International Organization for Standardization  
Business Owner: Director General Information -  
Business POC: CDMA10 - Rowan Edwards  
Existence Protective Marking: Unclassified  
Definition Protective Marking: Unclassified  
Originators Comment: This definition has already been approved as an OBJECT.  
Encyclopaedic POC: -  
Supporting Information URL: None  
Synonym: None  
Supporting Document: None  
**USAGE**  
Approved  
Usage Type: **Entity**  
DMG: Central Data Management Authority  
Additional Constraint:  
Comment:  
System Usage: used as DATA\_DEFINITION (Restricted) within [Defence Data Repository Application](#)  
Usage Link: has the [DATA DEFINITION APPROVAL STATUS](#) (Attribute - Approved)  
Usage Link: has the [DATA DEFINITION CENTRAL DATA MANAGEMENT INSTRUCTION](#) (Attribute - Approved)  
Usage Link: has the [DATA DEFINITION DEFINITION PROTECTIVE MARKING](#) (Attribute - Approved)  
Usage Link: has the [DATA DEFINITION DESCRIPTION](#) (Attribute - Approved)

# Data system infrastructure

1. Comprehensive – tabular, spatial, and hardcopy
2. Integrated – across discipline, time, and space
3. Efficient – not duplicative
4. Utilizes off-the-shelf products – Microsoft Windows, PC platform, ESRI, Oracle, Follet
5. Automated where ever possible – leveraging technology

# Documentation

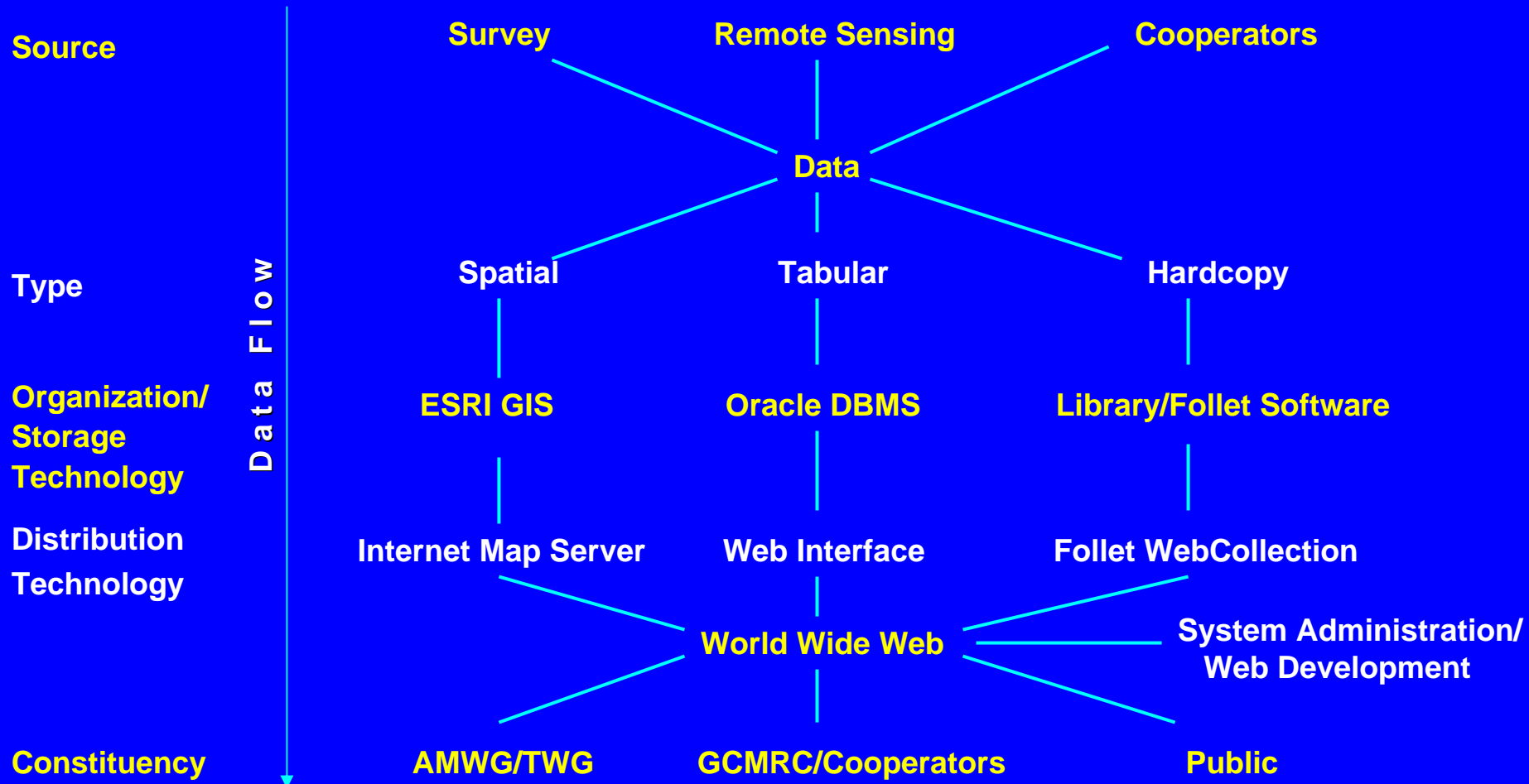
1. Design and structure
2. Troubleshooting
3. Maintenance
4. System backup and recovery
5. Metadata
6. Etc.

# GCMRC data management goals

1. Ensure data integrity
2. Consolidate all data
3. Enhances data longevity
4. Promotes data accessibility
5. Facilitate data analysis
6. Enable data integration
7. Metadata present and accessible
8. Provide data security
9. Stored in common formats



# Grand Canyon Monitoring and Research Center Data Infrastructure Diagram



# Why data management?

1. Data is our largest and most worthwhile asset
2. Data collection represents our largest cost
3. Data is cumulative if collected and archived properly
4. Current data management practices will determine the degree to which we are able to use contemporary data in the future
5. Good management decisions depend on good data
6. Facilitates Interoperability
7. Stimulates new thinking and fresh ideas
8. Helps manage complexity

# Data Management at the Grand Canyon Monitoring and Research Center

*by*

Michael J. Liszewski

Grand Canyon Monitoring and Research Center



# Why not data management?

1. It's boring
2. It's frustrating
3. It's inconvenient
4. It takes more time
5. It costs more money
6. It takes resources away from “science”