Data Management: The Hidden Enabler or (The Key Data and Work Product Integrator)

5th Annual CMMI Technology Conference & User Group 14-17 November 2005

Gary F. Norausky
&
Les Stamnas

Data Management, Configuration Management and the CMMI

- Data Management and Configuration Management according to CMMI:
  - ✓ CMMI V1.1's glossary contains the following definitions:
    - *Configuration Management:* A discipline applying technical and administrative direction and surveillance to (1) identify and document the functional and physical characteristics of a configuration item, (2) control changes to those characteristics, (3) record and report change processing and implementation status, and (4) verify compliance with specified requirements.
    - *Data Management:* Principles, processes, and systems for the sharing and management of data.
      - Project Planning Process Area (SG2, SP 2.3-1) and the monitoring of project data within the Project Monitoring And Control Process Area (SG1, SP 1.4-1).
      - Data is described in terms of “documentation,” and thus the confusion begins
      - Data Management consists of processes and systems that plan, acquire, and provide control for product and product-related business data, consistent with requirements, throughout the product and data life cycles.
  - ✓ Misconception: “Data” is somehow a “new thing,” if it is considered in the Project Planning process for management purposes

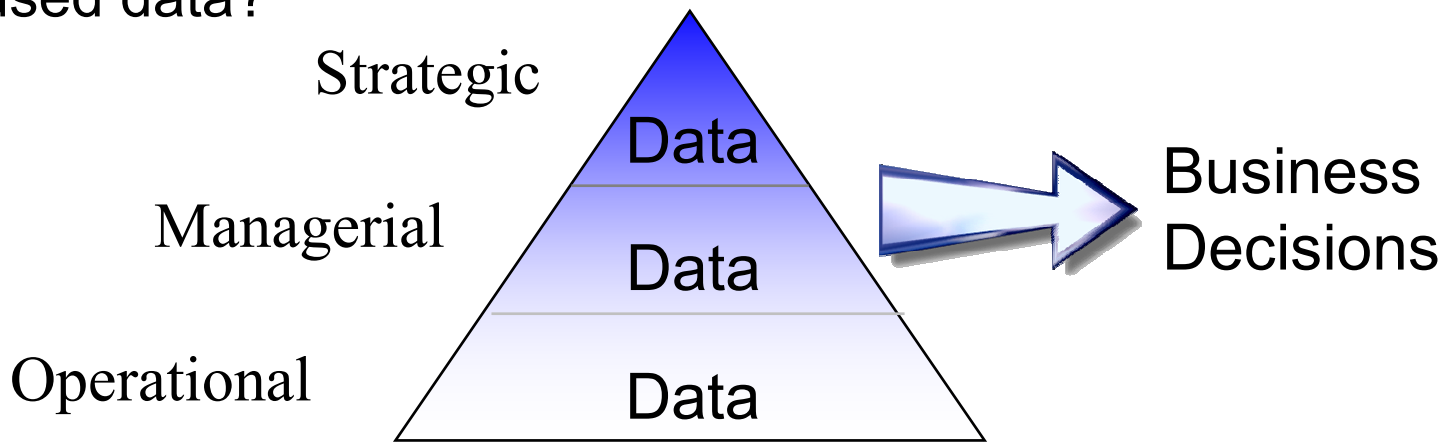
What is Data?

- Information in various forms
  - ✓ Managerial
  - ✓ Financial
  - ✓ Technical
  - ✓ Engineering
  - ✓ Administrative
  - ✓ Security
  - ✓ Procurement

NOTE: Data is essentially anything other than hardware, software and interfaces. It includes but not limited to cost and status reports, drawings, documents, source code, and listings, etc.

Why Data is Important?

- Useful predictions require an analysis of a lot of data - the more the better - and it should be relevant to your business and the environment in which you operate
- Data forms a significant and important element within data-driven systems, one would expect that the development methods used to produce it would reflect the same degree of care and attention that is applied to the other systems' components
- The key manager question is, "Where do I get relevant, updated, focused data?"



Key Data Management Considerations

General Requirements

Data Capabilities

Process and Product Quality Assurance

Data Processes

External Constraints

Related Business Objectives

Intended Use of the Data

Configuration Management

Technology Issues

Risk Management

Measurement and Analysis

Decision Analysis and Resolution

Data Management as a Functional Enabler

- Data Management's solution must address a functional need
- Data's value is not limited to its use in support of a particular product:
  - ✓ Data may have a life cycle longer than that of the product it describes, e.g., data from previous projects forms part of the foundation for new product and process design.
- Data also supports the enterprise in process redesign and quality.
  - ✓ Data is essential to competitive position.
  - ✓ Data is an integral part of an enterprise's intellectual assets and overall enterprise knowledge.
- Inaccurate or inconsistent data can hinder your company's ability to understand its current – and future business problems.

Data Management Plan

The prime functions of efficiency and effectiveness are:

- Administration of contract record keeping
- Data copying control
- Data quality control
- Acquisition/administration of supplier data
- Storage and retrieval systems
- Handling of classified data
- Maintenance and control of supplier-developed inform
  Purchaser-furnished information
- Identification and handling of property rights-in-data
- Pricing data
- Planning, scheduling, and delivery of data

30-50% of application design time is spent on copy management. Source: IBM

30% of people's time: searching for relevant information. Source: IBM

85% of information is unstructured. Source IBM

Data Risks

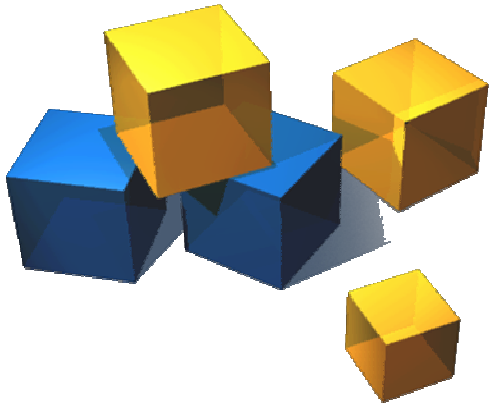
Data is often:

- Not subjected to any systematic hazard or risk analysis
- Poorly managed or controlled
- Not given any specific safety requirements
- Not assigned any specific integrity requirements
- Poorly structured, making errors more likely and harder to detect
- Not subjected to any form of verification
- Drawn from a single source

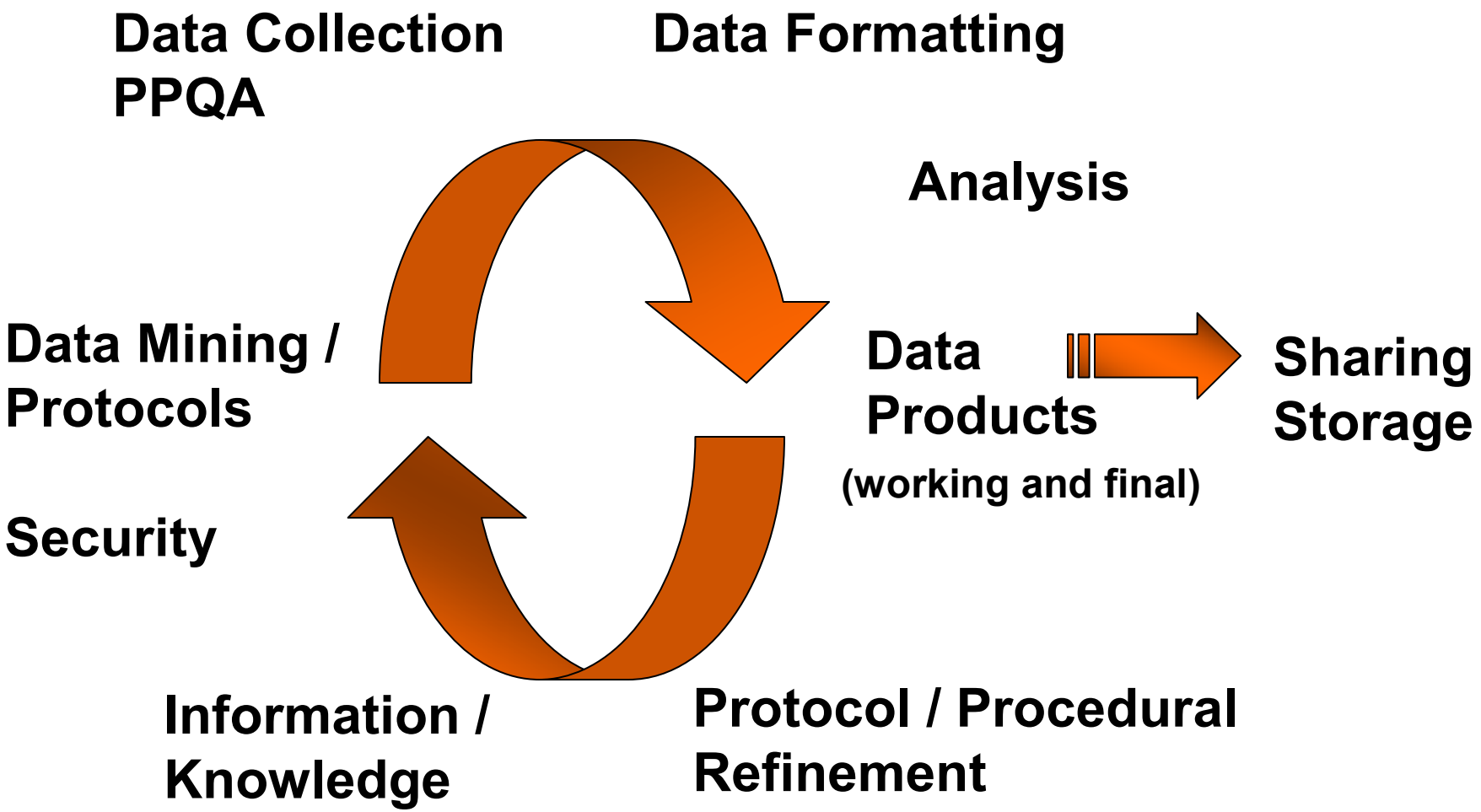
N. Storey and A. Faulkner, Data Management in Safety-Related Systems, *Proc. 20th International System Safety Conference*, Denver, 2002.

Generic Building Blocks for Data Management

- **Data Profiling** – Discover and analyze data discrepancies
- **Data Quality** – Reconcile and correct data and improve the processes that create it
- **Data Integration** – Integrate and link data across disparate sources
- **Data Augmentation** – Enhance information using internal or external data sources
- **Data Monitoring** – Check and control data integrity over time
- **Knowledge Management** – Ensure data is accurate and that the filters, relations and criteria are captured to provide context for the information reserve



Data Lifecycle and Work Flow



Steps for Developing a Data Design

- Step 1 - Preliminary scoping meetings – data managers and protocol developers/project leaders
  - Review the context, purpose and sources of project data
  - Clarify how data are acquired, entered, processed and documented
  - Who performs these steps, and the quality control measures built into these processes?
  - Discuss the timing and frequency of data entry and updates
  - Who needs access to the data at different stages of the data life cycle
  - Could certain project data qualify for protection as sensitive info?
  - What are the project needs for data distribution? When and how should data be made available to others?

Steps for Developing a Data Design

- ✓ Step 2 – Develop the “logical” data model (tables, fields, data types, domains, range limits, descriptions)
- ✓ Step 3 – Have this reviewed to make sure it meets network and national standards, and fits the project needs
- ✓ Step 4 – Complete the data design to address specific implementation details
  - Define in detail the integration needs with other past, present or future data sets
  - Identify and define needed data views. How does information need to be summarized, presented and exported? How do geographic data need to be displayed?
  - What is the intended audience for different products, and what are their specific needs?

Steps for Developing a Data Design

- Step 4 – continued …
  - Where will working data reside?
  - What will the software platform be for database implementation?
  - Does the project require a separate working database for current year data? If so, what milestone(s) must occur prior to data being uploaded into the master database for summarization and analysis?
  - Identify the types of data backups that might need to be made, and the specific project milestones that trigger these backups
  - Specify how and when certified data sets will be delivered
  - Define measures & responsibilities for protecting sensitive information
  - Clarify responsibilities and expectations for database maintenance. Are there sufficient resources to maintain the database as it is scoped?

Steps for Developing a Data Design

- ✓ Step 5 – Develop the “physical” database model (i.e., create the database)
- ✓ Step 6 – Develop the application interface - data entry, processing, summarization and reporting, exports for analysis
- ✓ Step 7 – Have everything tested and reviewed to make sure it works and meets project needs
- ✓ Step 8 – Develop documentation and training materials

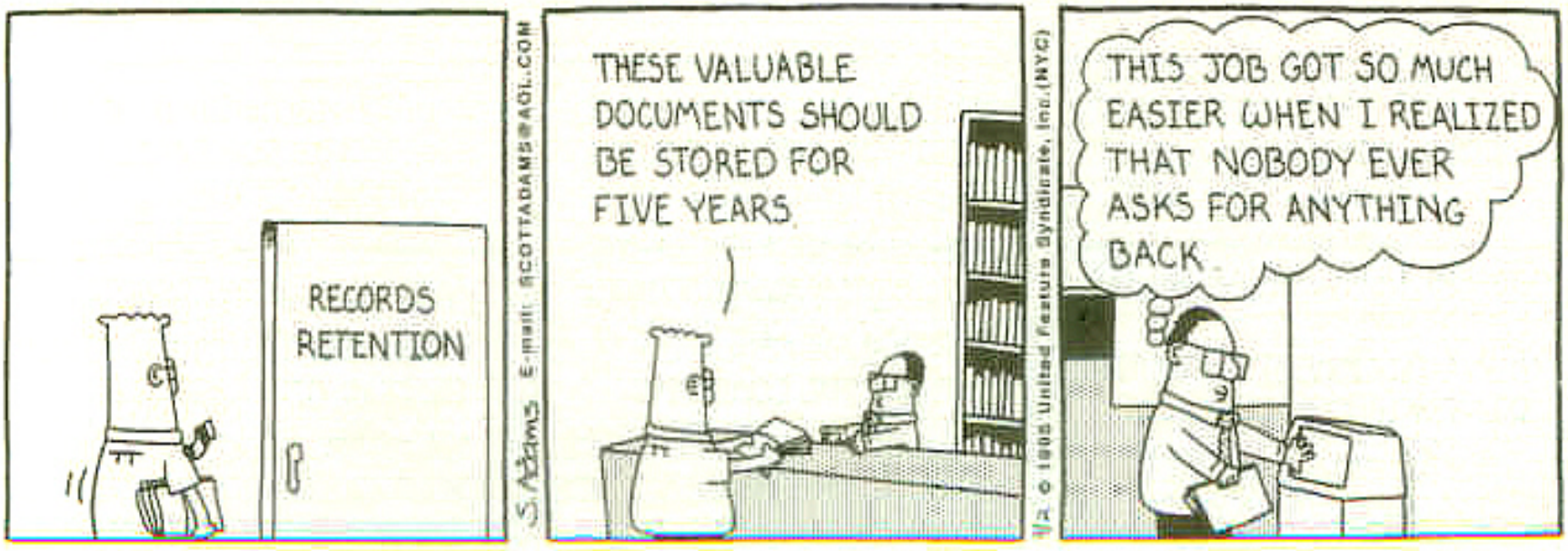
Example Standard for Data Management

- Government Electronics and Information Technology Association (GEIA) 859-2004
  - ✓ Describes DM principles and methods using a neutral DM terminology.
  - ✓ Intended to articulate contemporary DM principles and methods that are broadly applicable to management of electronic and non-electronic data in both the commercial and government sectors.
  - ✓ Addresses product data and the business data intrinsic to collaboration during product acquisition and sustainment.

GEIA 859 Data Management Principles

ID	Area	Principle
1	Focus and Scope	Define the organizationally-relevant scope relevant scope of data management.
2	Customer Support	Plan for, acquire, and provide data responsive to customer requirements.
3	Business Context	Develop DM processes to fit the context and business environment in which they will be performed.
4	Identification	Identify data products and views so that their requirements and attributes can be controlled.
5	Change Management	Control data, repositories, data products, data views, and metadata using approved change control processes.
6	Data Rights	Establish and maintain an identification,process for intellectual property, proprietary, and competition-sensitive data.
7	Data Retention	Retain data commensurate with value to the organization.
8	Process Improvement	Continuously improve data management.
9	DM/KM Connection	Effectively integrate data management (DM) with knowledge management (KM).

Not A Good Data Management Repository



Contact Us

Gary F. Norausky
norausky@norauskypsi.com
+1(619) 472 8810

Les Stamnas
stamnas@norauskypsi.com
+1(858) 735 3965

Norausky Process Solutions, Inc.
www.norauskypsi.com