MIS 5206
Protecting Information Assets
- Unit# 2b -

Data Classification Processes and Models

Agenda

- Vocabulary
- Data Classification Process and Models
- Test taking tip
- Quiz

Information Systems Security Controls

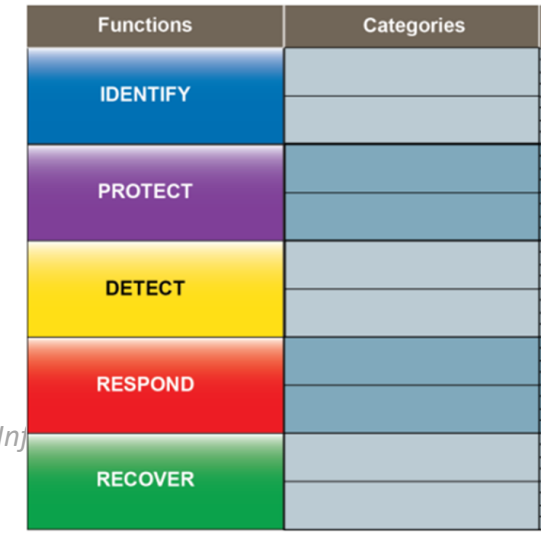
TABLE 1: SECURITY CONTROL IDENTIFIERS AND FAMILY NAMES

ID	FAMILY	ID	FAMILY
AC	Access Control	MP	Media Protection
AT	Awareness and Training	PE	Physical and Environmental Protection
AU	Audit and Accountability	PL	Planning
CA	Security Assessment and Authorization	PS	Personnel Security
CM	Configuration Management	RA	Risk Assessment
CP	Contingency Planning	SA	System and Services Acquisition
IA	Identification and Authentication	SC	System and Communications Protection
IR	Incident Response	SI	System and Information Integrity
MA	Maintenance	PM	Program Management

Taxonomies of InfoSys Controls

By Function

- Identify
- Protect
- Detect
- Respond
- Recover



By Class

- Management
- Operational
- Technical

CLASS	FAMILY	IDENTIFIER
Management	Risk Assessment	RA
Management	Planning	PL
Management	System and Services Acquisition	SA
Management	Certification, Accreditation, and Security Assessments	CA
Operational	Personnel Security	PS
Operational	Physical and Environmental Protection	PE
Operational	Contingency Planning	CP
Operational	Configuration Management	CM
Operational	Maintenance	MA
Operational	System and Information Integrity	SI
Operational	Media Protection	MP
Operational	Incident Response	IR
Operational	Awareness and Training	AT
Technical	Identification and Authentication	IA
Technical	Access Control	AC
Technical	Audit and Accountability	AU
Technical	System and Communications Protection	SC

Taxonomies of InfoSys Controls

By <u>Modality</u>

1. Physical
2. Technical
3. Administrative

A modality is the way (or mode) in which something is done

http://www.sans.edu/research/security-laboratory/article/security-controls

Taxonomies of InfoSys Controls

By Phase

1. Preventative
2. Detective
3. Corrective

Preventative	Detective	Corrective	Compensatory
Security Awareness Training	System Monitoring	OS Upgrade	Backup Generator
Firewall	IDS	Backup Data Restoral	Hot Site
Anti-virus	Anti-Virus	Anti-Virus	Server Isolation
Security Guard	Motion Detector	Vulnerability Mitigation	
IPS	IPS		

These are sometimes referred to as "*phase controls*"

http://www.sans.edu/research/security-laboratory/article/security-controls

Taxonomies of InfoSys Controls

By function

- Preventive
- Detective
- Corrective
- Compensating

By modality

- Physical
- Technical
- Administrative

Juxtaposing taxonomies to improve understanding…

Modality

Function	Controls	Administrative	Technical	Physical
	Preventive	*User registration*	*Passwords, Tokens*	*Fences*
	Detective	*Report reviews*	*Audit Logs*	*Sensors*
	Corrective	*Employee termination*	*Connection management*	*Fire extinguisher*
	Compensating	*Supervision*	*Keystroke logging*	*Layered defenses*

Question

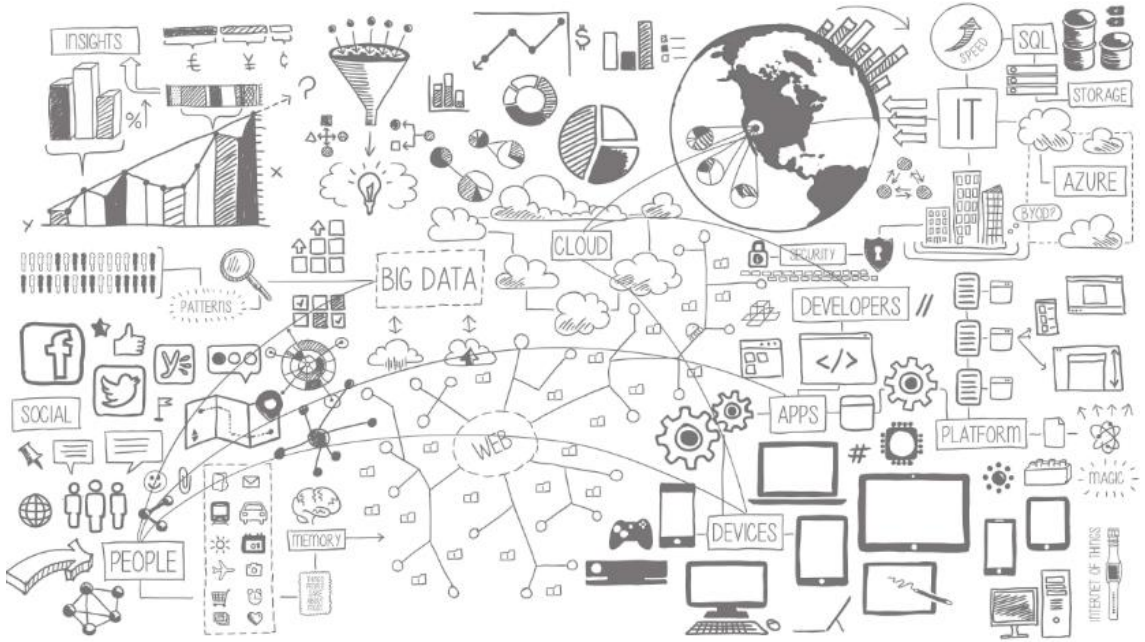
- What is data ?
- What is information ?
- How do data and information relate to each other?
- What is an information system?

What is data ?

http://researchdata.ox.ac.uk/

1. Known facts or things used as a basis for inference or reckoning
2. Quantities or characters operated on by a computer etc.

The Concise Oxford Dictionary



https://blogs.microsoft.com/blog/2014/04/15/a-data-culture-for-everyone/

What is the nature of data stored in the attributes comprising the entities within the information system's databases

What is information?

*An Entity's attribute values can be understood in terms of **"measurement levels"***

Stevens, S.S. 1946. On the theory of scales of measurement. Science 103:677-680.

Measurements levels describe the inherent nature of information in the attribute data that make up entities

- Qualitative information tells what things exist
- Quantitative information orders and measures the magnitude of these things

Steven's 4 measurement levels

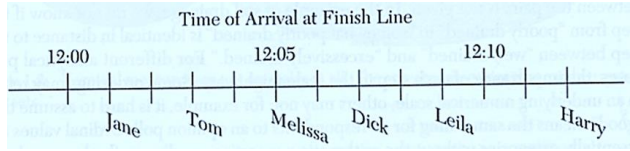
1. Nominal
2. Ordinal
3. Interval
4. Ratio

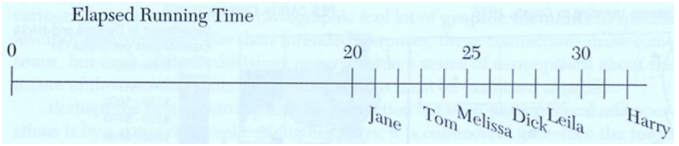
Increasing information content

Measurement Levels

Scale	Defining Relations
Nominal	(a) Equivalence Class A = Class A Class A ≠ Class B
Ordinal	(a) Equivalence (b) Greater-less than A > B B < A
Interval	(a) Equivalence (b) Greater-less than
Ratio	(a) Equivalence (b) Greater-less than (c) Ratio of any two intervals (d) Ratio of any two scale values (assumed true 0 value)

Order of arrival of contestants	*Women's race*	*Men's race*
First	Jane	Tom
Second	Melissa	Dick
Third	Leila	Harry

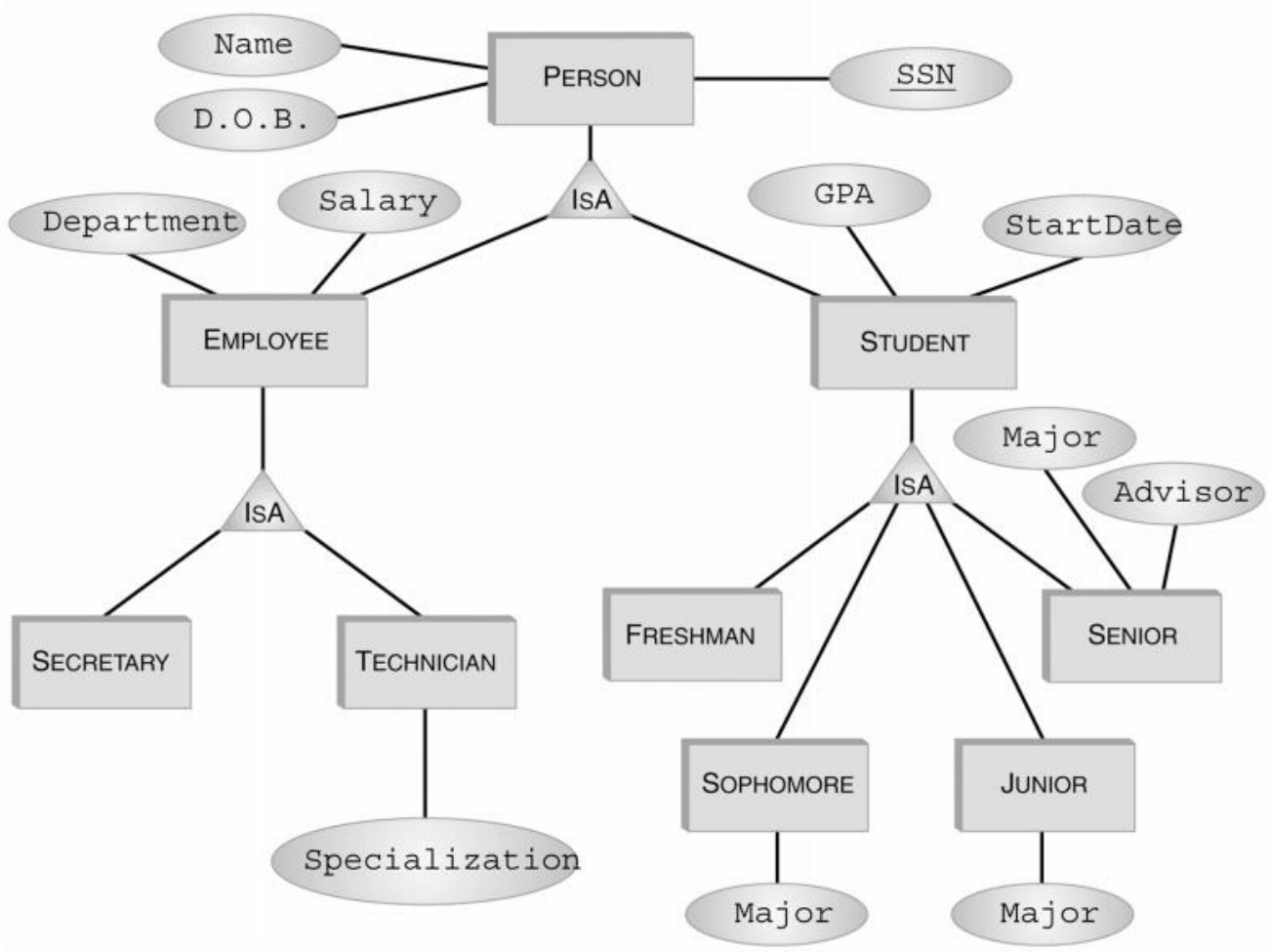




Entity Attribute Value Measurement Types

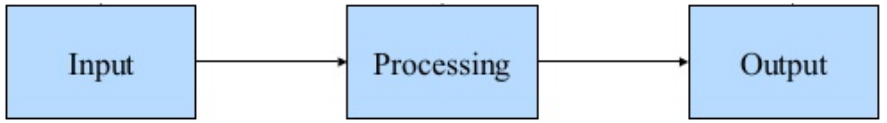
	Qualitative	**Quantitative**
Nominal	X	
Ordinal	X	
Interval		X
Ratio		X

How would you use Steven's measurements levels to categorize this information ?

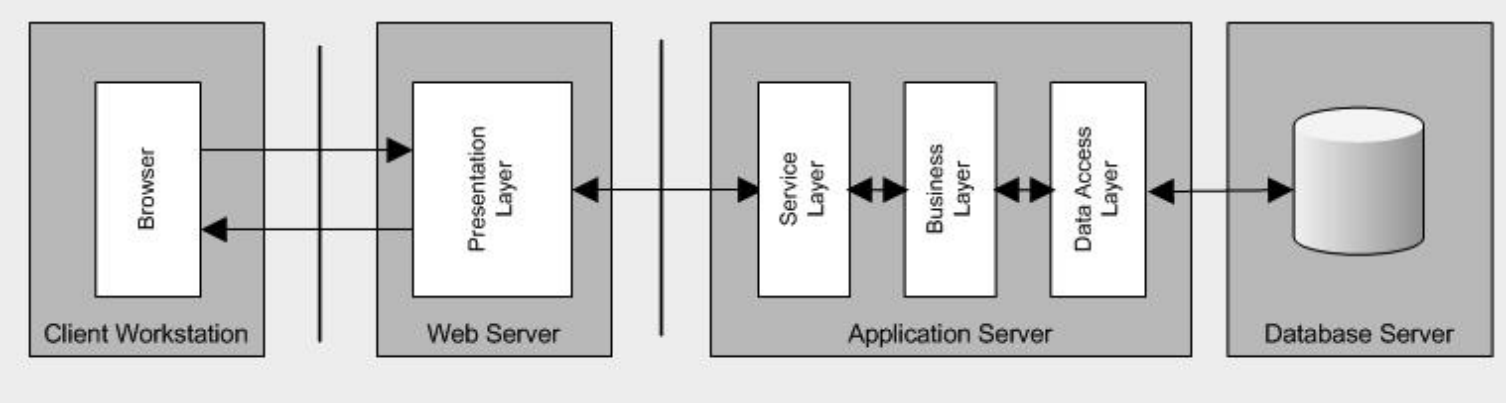


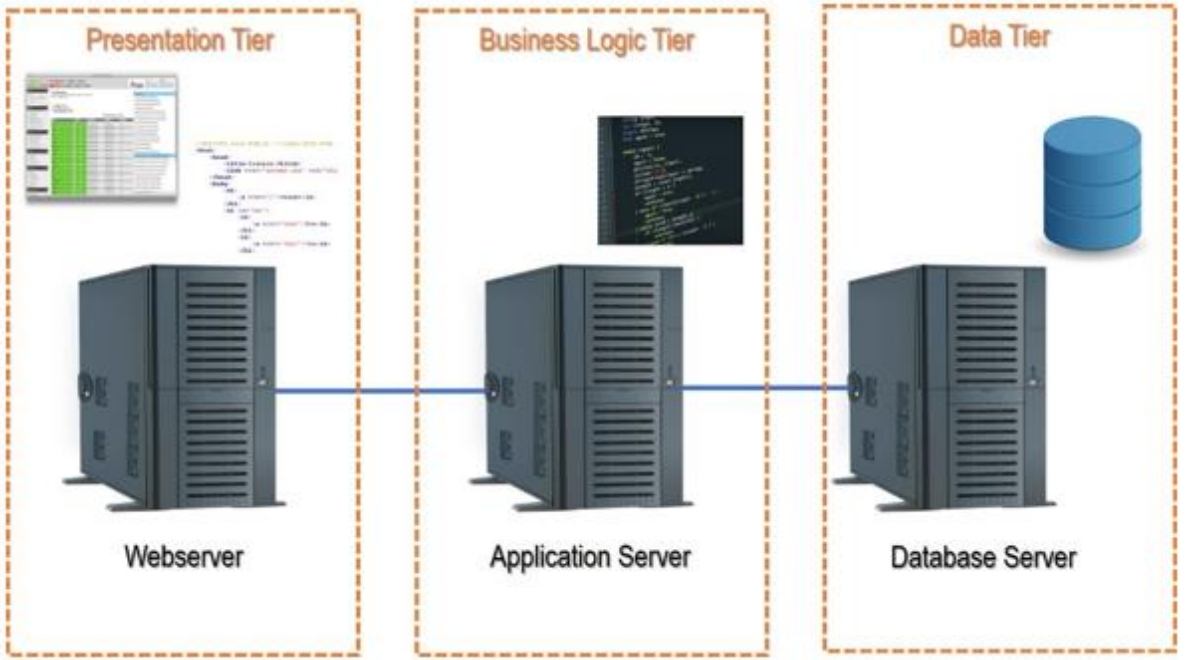
What is an information system

"An **information system** (**IS**) is an organized system for the collection, organization, storage and communication of information. …complementary networks that people and organizations use to collect, filter (query), process, create and distribute data. Further, an information system (IS) is a group of components that interact to produce information." Wikepedia

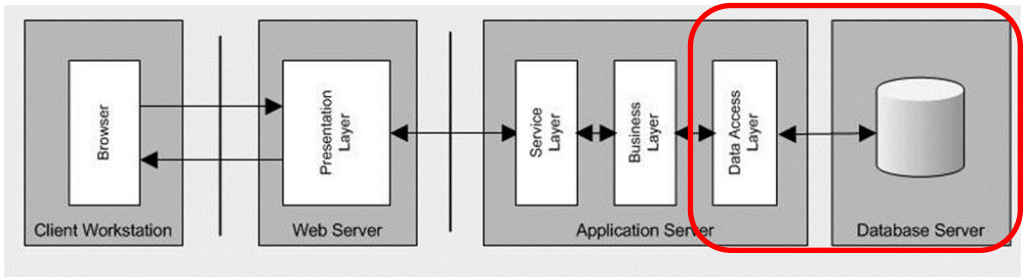


Information system (IS) architectures

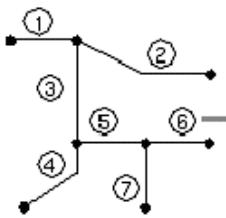




Information System Data



Coverage: Roads



Roads #	x,y Coordinates
1	2,12 6,12
2	6,12 10,10 14,10
3	6,6 6,12
4	3,2 6,4 6,6
5	6,6 10,6
6	10,6 14,6
7	10,2 10,6

Road Number	Road Type	Surface	Width	Lanes	Name
1	1	Concrete	60	4	Hwy 42
2	1	Concrete	60	4	Hwy 42
3	2	Asphalt	48	4	N Main St.
4	2	Asphalt	48	4	N Main St.
5	3	Asphalt	32	2	Cedar Ave.
6	3	Asphalt	32	2	Cedar Ave.
7	4	Asphalt	32	2	Elm St.

Relational Data Model

Sid #	Name	Year	GPA
1	Smith	3	3.0
2	Jones	2	3.5
3	Doe	1	1.2
4	Varda	4	4.0
5	Carey	4	0.5

Student Relation

Fid #	Name	Position	Dept
9	Henry	Prof.	Math
2	Jackson	Assist. Prof	Hist
14	Schuh	Assoc. Prof	Chem
21	Lerner	Assist. Prof	CS

Faculty Relation

C #	Course Name	Cr	Dept
223	Calculus	5	Math
302	Intro Prog	3	CS
302	Organic Chem	3	Chem
542	Asian Hist	2	Hist
222	Calculus	5	Math

Course Relation

Taught-By Relation

C #	Fid #
223	9
222	9
302	21
302	14
542	2

Enrolled Relation

Sid #	C #
1	223
4	222
4	302
3	302
5	302
2	542
2	223

Concept

Classification Grouping of data according to pre-determined types

Why classify data ?

Data Classification Processes and Models

Data classification ("categorization") is essential to ensuring that data is appropriately protected, and done so in the most cost-effective manner

The goal is to classify data according to risk associated with a breach to their confidentiality, integrity, and availability

Enables determining the appropriate cost expenditure of security control mitigations required to protect the IT assets

Key Concepts

Classification	Grouping of data according to pre-determined types
Cost-Effectiveness	Appropriateness of the level of risk mitigation expenditure
Confidentiality	Restriction who may know about and/or have access to information
Integrity	Confidence that information is complete and unaltered
Availability	Access to information

Question:

How should we determine the information security categorization of an IT asset?

FIPS 199 Standards: Security objectives and impact ratings

FIPS PUB 199

FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION

Standards for Security Categorization of Federal Information and Information Systems

Low: *Limited adverse effect*

Moderate: *Serious adverse effect*

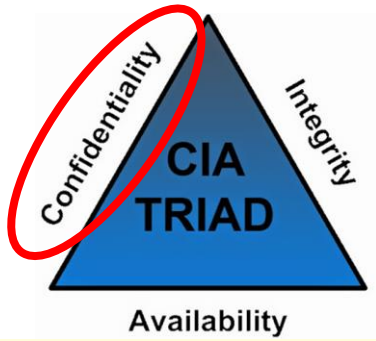
High: *Severe or catastrophic adverse effect*

	POTENTIAL IMPACT		
Security Objective	**LOW**	**MODERATE**	**HIGH**
Confidentiality Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. [44 U.S.C., SEC. 3542]	The unauthorized disclosure of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Integrity Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity. [44 U.S.C., SEC. 3542]	The unauthorized modification or destruction of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Availability Ensuring timely and reliable access to and use of information. [44 U.S.C., SEC. 3542]	The disruption of access to or use of information or an information system could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

FIPS PUB 199

FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION

Standards for Security Categorization of Federal Information and Information Systems

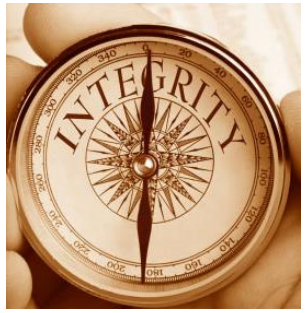


	POTENTIAL IMPACT		
Security Objective	**LOW**	**MODERATE**	**HIGH**
Confidentiality Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. [44 U.S.C., SEC. 3542]	The unauthorized disclosure of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

FIPS PUB 199

FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION

Standards for Security Categorization of Federal Information and Information Systems



	POTENTIAL IMPACT		
Security Objective	**LOW**	**MODERATE**	**HIGH**
Integrity Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity. [44 U.S.C., SEC. 3542]	The unauthorized modification or destruction of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

FIPS PUB 199

FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION

Standards for Security Categorization of Federal Information and Information Systems

	POTENTIAL IMPACT		
Security Objective	**LOW**	**MODERATE**	**HIGH**
Availability Ensuring timely and reliable access to and use of information. [44 U.S.C., SEC. 3542]	The disruption of access to or use of information or an information system could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

FIPS 199 standard: Security objectives and impact ratings

Low: *Limited adverse effect*

Moderate: *Serious adverse effect*

High: *Severe or catastrophic adverse effect*

What kind of Steven's measurement level is used by the FIPS 199 Information Security categorization standard?

	POTENTIAL IMPACT		
Security Objective	**LOW**	**MODERATE**	**HIGH**
Confidentiality Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. [44 U.S.C., SEC. 3542]	The unauthorized disclosure of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Integrity Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity. [44 U.S.C., SEC. 3542]	The unauthorized modification or destruction of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Availability Ensuring timely and reliable access to and use of information. [44 U.S.C., SEC. 3542]	The disruption of access to or use of information or an information system could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

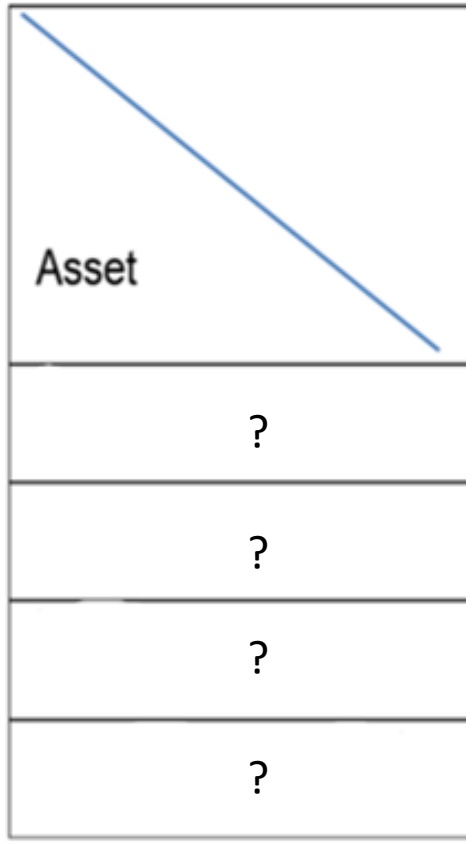
Question:

How do you determine the information security categorization of each dataset on the Dean's computer?

Exercise:

1. *Inventory the (possible) types of information that might be on the Dean's laptop*
2. *Assign information security categorizations to the information contained on the Dean's laptop*
3. *Provide an overall security categorization for the laptop*

1. Create an inventory of types of datasets possibly stored on the Dean's laptop



2. Assign information security categorization impact ratings to the data on the Dean's laptop…

Impact to / Asset	Confidentiality	Integrity	Availability
Staff Salary Data			
Student Data			
Fundraising Presentations			
Dean's Personal Data			

What is the FIPS 199 information security categorization of the Dean's laptop?

Impact to / Asset	Confidentiality	Integrity	Availability
Staff Salary Data	High	Low	Medium
Student Data	High	Low	Low
Fundraising Presentations	Medium	Medium	High
Dean's Personal Data	Low	Low	Medium
Overall Impact	?	?	?

FIPS Pub 199 Standard for determining the security categorization of an information system that contains or transports multiple information types

The generalized format for expressing the security category, SC, of an information system is:

$$\text{SC}_{\text{information system}} = \{(\textbf{confidentiality}, \textit{impact}), (\textbf{integrity}, \textit{impact}), (\textbf{availability}, \textit{impact})\},$$

where the acceptable values for potential impact are LOW, MODERATE, or HIGH.

Example with multiple information types:

$$\text{SC}_{\text{contract information}} = \{(\textbf{confidentiality}, \text{MODERATE}), (\textbf{integrity}, \text{MODERATE}), (\textbf{availability}, \text{LOW})\},$$

and

$$\text{SC}_{\text{administrative information}} = \{(\textbf{confidentiality}, \text{LOW}), (\textbf{integrity}, \text{LOW}), (\textbf{availability}, \text{LOW})\}.$$

The resulting security category of the information system is expressed as:

$$\text{SC}_{\text{acquisition system}} = \{(\textbf{confidentiality}, \text{MODERATE}), (\textbf{integrity}, \text{MODERATE}), (\textbf{availability}, \text{LOW})\},$$

Low: Limited adverse effect
Moderate: Serious adverse effect
High: Severe or catastrophic adverse effect

Overall impact in each of the CIA dimensions is based on the highest impact dataset in each of the dimensions

Impact to / Asset	Confidentiality	Integrity	Availability
Staff Salary Data	High	Low	Medium
Student Data	High	Low	Low
Fundraising Presentations	Medium	Medium	High
Dean's Personal Data	Low	Low	Medium
Overall Impact	**High**	**Medium**	**High**

What single overall information security categorization would you give each dataset on the Dean's laptop?

Impact to / Asset	Confidentiality	Integrity	Availability	Categorization
Staff Salary Data	High	Low	Medium	?
Student Data	High	Low	Low	?
Fundraising Presentations	Medium	Medium	High	?
Dean's Personal Data	Low	Low	Medium	?
Overall Impact	High	Medium	High	

Single overall information security impact ratings for each dataset on the Dean's laptop

Impact to / Asset	Confidentiality	Integrity	Availability	Categorization
Staff Salary Data	High	Low	Medium	**High**
Student Data	High	Low	Low	**High**
Fundraising Presentations	Medium	Medium	High	**High**
Dean's Personal Data	Low	Low	Medium	**Medium**
Overall Impact	High	Medium	High	

What single value would you use to rate the information security requirements of the Dean's laptop?

Impact to / Asset	Confidentiality	Integrity	Availability	**Categorization**
Staff Salary Data	High	Low	Medium	High
Student Data	High	Low	Low	High
Fundraising Presentations	Medium	Medium	High	High
Dean's Personal Data	Low	Low	Medium	Medium
Overall Impact	High	Medium	High	?

The single overall information security categorizations for each dataset on the Dean's laptop

Impact to / Asset	Confidentiality	Integrity	Availability	**Categorization**
Staff Salary Data	High	Low	Medium	High
Student Data	High	Low	Low	High
Fundraising Presentations	Medium	Medium	High	High
Dean's Personal Data	Low	Low	Medium	Medium
Overall Impact	High	Medium	High	**High**

How do you define and relate the following to each other?

- Policy
- Standard
- Guideline
- Procedure

Policy, Standard, Guideline and Procedures

- **Policy:** A formal, brief, and high-level statement or plan that embraces an organization's general beliefs, goals, objectives, and acceptable procedures for a specified subject area. Policies always state required actions, and may include pointers to standards. Policy attributes include the following:
  - Requires compliance (mandatory)
  - Failure to comply results in disciplinary action
  - Focus on desired results, not on means of implementation
  - Further defined by standards and guidelines

- **Standard:** A mandatory action or rule designed to support and conform to a policy.
  - A standard should make a policy more meaningful and effective.
  - A standard must include one or more accepted specifications for hardware, software, or behavior.

- **Guideline:** General statements, recommendations, or administrative instructions designed to achieve the policy's objectives by providing a framework within which to implement procedures.
  - A guideline can change frequently based on the environment and should be reviewed more frequently than standards and policies.
  - A guideline is not mandatory, rather a suggestion of a best practice. Hence "guidelines" and "best practice" are interchangeable

- **Procedures:** Procedures describe the process: who does what, when they do it, and under what criteria. They can be text based or outlined in a process map.
  - A series of steps taken to accomplish an end goal.
  - Procedures define "how" to protect resources and are the mechanisms to enforce policy.
  - Procedures provide a quick reference in times of crisis.
  - Procedures help eliminate the problem of a single point of failure.
  - Also known as a SOP (Standard Operating Procedure)

Policy Example

RA-2	SECURITY CATEGORIZATION	
	ASSESSMENT OBJECTIVE: *Determine if the organization:*	
	RA-2(a)	*categorizes information and the information system in accordance with applicable federal laws, Executive Orders, directives, policies, regulations, standards, and guidance;*
	RA-2(b)	*documents the security categorization results (including supporting rationale) in the security plan for the information system; and*
	RA-2(c)	*ensures the authorizing official or authorizing official designated representative reviews and approves the security categorization decision.*
	POTENTIAL ASSESSMENT METHODS AND OBJECTS: **Examine**: [*SELECT FROM:* Risk assessment policy; security planning policy and procedures; procedures addressing security categorization of organizational information and information systems; security plan; security categorization documentation; other relevant documents or records]. **Interview**: [*SELECT FROM:* Organizational personnel with security categorization and risk assessment responsibilities; organizational personnel with information security responsibilities]. **Test**: [*SELECT FROM:* Organizational processes for security categorization].	

Information Technology & Telecommunications

The City of New York

CITYWIDE INFORMATION SECURITY POLICY

Data Classification Policy

The Policy

The Agency head or designee has responsibility for ensuring agency information assets are appropriately categorized and the appropriate degree of protection is applied based on its valuation.

Background

To ensure that business information assets receive an appropriate level of protection, the value of the information must be assessed to determine the requirements for security protection. Business information assets are those that affect and are integral to the City's ability to provide business services with integrity, comply with laws and regulations, and meet public trust.

Scope

This policy applies to all information. Information is defined as anything spoken, overheard, written, stored electronically, copied, transmitted or held intellectually concerning the City of New York general business, information systems, employees, business partners, or customers.

Information Classification

All information at the City of New York and corresponding agencies will be classified at one of four levels; public, sensitive, private, or confidential.

- **Public**—This information might not need to be disclosed, but if it is, it shouldn't cause any damage.
- **Sensitive**—This information requires a greater level of protection to prevent loss of inappropriate disclosure.
- **Private**—This information is for agency use only, and its disclosure would damage the public trust placed in the agency.
- **Confidential**—This is the highest level of sensitivity, and disclosure could cause extreme damage to the agency's ability to perform its primary business function. Datasets containing information whose disclosure could lead directly to massive financial loss, danger to public safety, or lead to loss of life is classified as confidential.

Information Valuation and Categorization

1) Ensure that business information assets receive an appropriate level of protection. The value of the information must be assessed to determine the requirements for security protection.
2) All information assets must be valued and categorized.
3) Information assets must be evaluated, valued and categorized by the Data Steward on a regular basis.
4) To ensure that appropriate protection is provided, the value of information should be determined before transmission over any communications network.

There are alternative “code word” classification systems for data

Clearance		
Highest	Top secret	Unauthorized disclosure could be expected to cause “exceptionally grave damage” to national security
2nd highest	Secret	Unauthorized disclosure would cause “serious damage” to national security
Lowest	Confidential	Unauthorized disclosure would “damage” national security
Unclassified	*Sensitive But Classified*	*Synonym*
Unclassified	*For Official Use Only (FOUO)*	*Synonym*
Unclassified	*Controlled Unclassified Information (CUI)*	*Synonym*
Unclassified	…	…

- *The need to protect information is different in different sorts of organizations*
- *A principal distinction exists between government and business*
- *But, the terms used for CUI overlap in scope and are often intermingled often resulting in confusion*

Which do you prefer?

FIPS 199 Standard

The generalized format for expressing the security category, SC, of an information system is:

SC information system = {(**confidentiality**, *impact*), (**integrity**, *impact*), (**availability**, *impact*)},

where the acceptable values for potential impact are LOW, MODERATE, or HIGH.

...or...

New York City Data Classification Policy

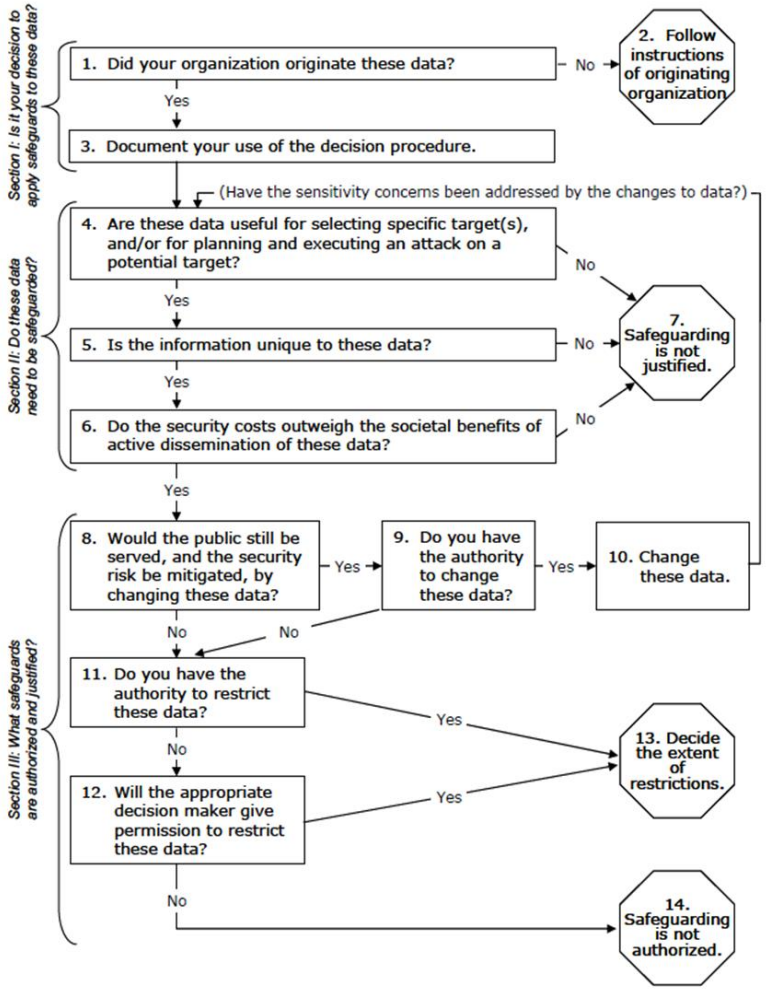
Information Classification

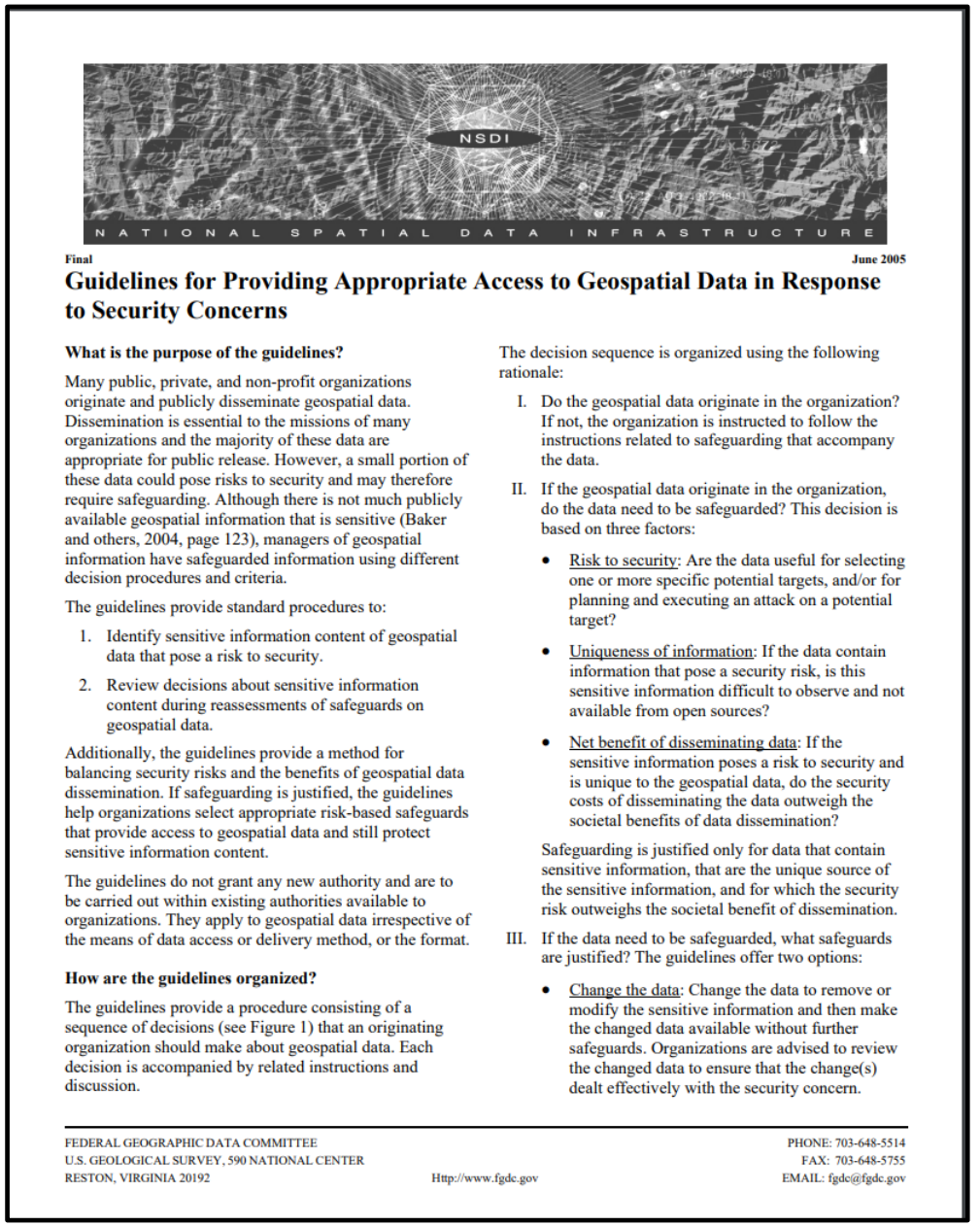
All information at the City of New York and corresponding agencies will be classified at one of four levels; public, sensitive, private, or confidential.

- **Public**—This information might not need to be disclosed, but if it is, it shouldn't cause any damage.
- **Sensitive**—This information requires a greater level of protection to prevent loss of inappropriate disclosure.
- **Private**—This information is for agency use only, and its disclosure would damage the public trust placed in the agency.
- **Confidential**—This is the highest level of sensitivity, and disclosure could cause extreme damage to the agency's ability to perform its primary business function. Datasets containing information whose disclosure could lead directly to massive financial loss, danger to public safety, or lead to loss of life is classified as confidential.

Why?

Analyzing datasets based on the need for confidentiality





Final June 2005

Guidelines for Providing Appropriate Access to Geospatial Data in Response to Security Concerns

What is the purpose of the guidelines?

Many public, private, and non-profit organizations originate and publicly disseminate geospatial data. Dissemination is essential to the missions of many organizations and the majority of these data are appropriate for public release. However, a small portion of these data could pose risks to security and may therefore require safeguarding. Although there is not much publicly available geospatial information that is sensitive (Baker and others, 2004, page 123), managers of geospatial information have safeguarded information using different decision procedures and criteria.

The guidelines provide standard procedures to:

1. Identify sensitive information content of geospatial data that pose a risk to security.
2. Review decisions about sensitive information content during reassessments of safeguards on geospatial data.

Additionally, the guidelines provide a method for balancing security risks and the benefits of geospatial data dissemination. If safeguarding is justified, the guidelines help organizations select appropriate risk-based safeguards that provide access to geospatial data and still protect sensitive information content.

The guidelines do not grant any new authority and are to be carried out within existing authorities available to organizations. They apply to geospatial data irrespective of the means of data access or delivery method, or the format.

How are the guidelines organized?

The guidelines provide a procedure consisting of a sequence of decisions (see Figure 1) that an originating organization should make about geospatial data. Each decision is accompanied by related instructions and discussion.

The decision sequence is organized using the following rationale:

I. Do the geospatial data originate in the organization? If not, the organization is instructed to follow the instructions related to safeguarding that accompany the data.

II. If the geospatial data originate in the organization, do the data need to be safeguarded? This decision is based on three factors:

- Risk to security: Are the data useful for selecting one or more specific potential targets, and/or for planning and executing an attack on a potential target?
- Uniqueness of information: If the data contain information that pose a security risk, is this sensitive information difficult to observe and not available from open sources?
- Net benefit of disseminating data: If the sensitive information poses a risk to security and is unique to the geospatial data, do the security costs of disseminating the data outweigh the societal benefits of data dissemination?

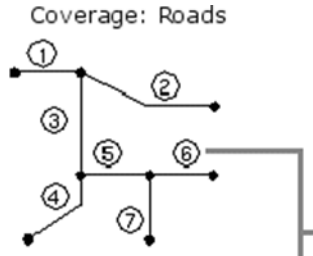
Safeguarding is justified only for data that contain sensitive information, that are the unique source of the sensitive information, and for which the security risk outweighs the societal benefit of dissemination.

III. If the data need to be safeguarded, what safeguards are justified? The guidelines offer two options:

- Change the data: Change the data to remove or modify the sensitive information and then make the changed data available without further safeguards. Organizations are advised to review the changed data to ensure that the change(s) dealt effectively with the security concern.

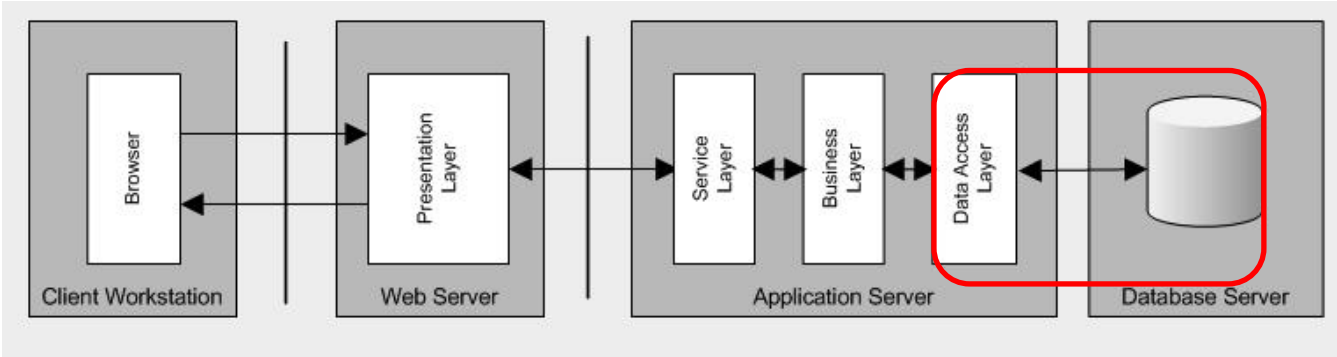
FEDERAL GEOGRAPHIC DATA COMMITTEE
U.S. GEOLOGICAL SURVEY, 590 NATIONAL CENTER
RESTON, VIRGINIA 20192
Http://www.fgdc.gov
PHONE: 703-648-5514
FAX: 703-648-5755
EMAIL: fgdc@fgdc.gov

Geo-Relational datasets



Roads #	x,y Coordinates
1	2,12 6,12
2	6,12 10,10 14,10
3	6,6 6,12
4	3,2 6,4 6,6
5	6,6 10,6
6	10,6 14,6
7	10,2 10,6

Road Number	Road Type	Surface	Width	Lanes	Name
1	1	Concrete	60	4	Hwy 42
2	1	Concrete	60	4	Hwy 42
3	2	Asphalt	48	4	N Main St.
4	2	Asphalt	48	4	N Main St.
5	3	Asphalt	32	2	Cedar Ave.
6	3	Asphalt	32	2	Cedar Ave.
7	4	Asphalt	32	2	Elm St.



Confidentiality categorization example…

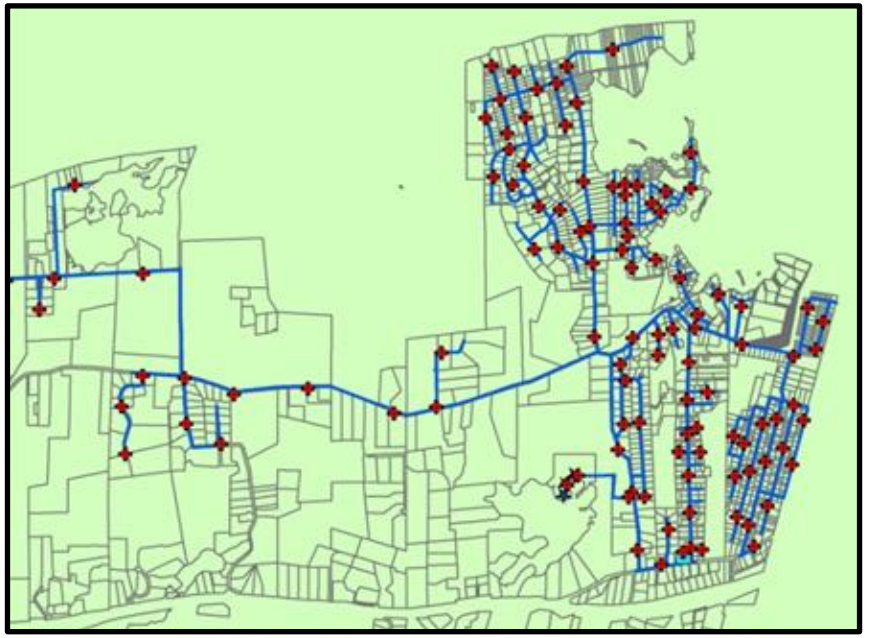
Framework for Analyzing the Homeland Security Sensitivity of Geospatial Data and Information Sources

Filter	Key Questions for Decisionmakers
Usefulness	• Is the information useful for target selection or location purposes? • Is the information useful for attack planning purposes?
Uniqueness	• Is the information readily available from other geospatial information sources? • Is the information available from direct observation or other nongeospatial information types?
Societal benefits and costs	• What are the expected security benefits of restricting public access to the source? • What are the expected societal costs of restricting public access to the source?

4. Are these data useful for selecting specific target(s), and/or for planning and executing an attack on a potential target?

Do the data show choke points" to increase effectiveness of an attack ?

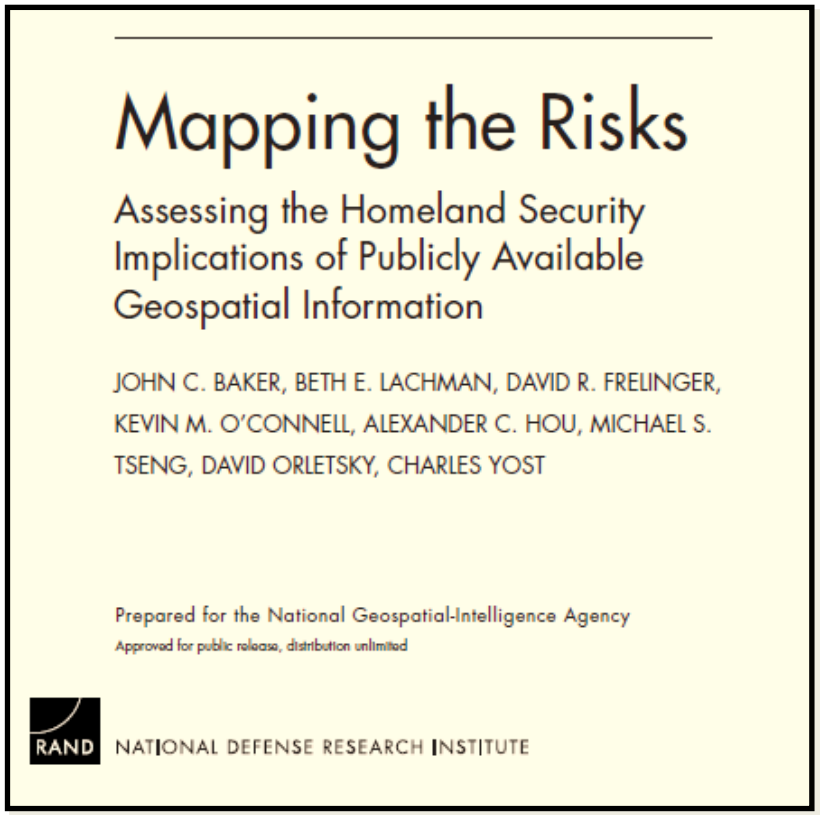
Do the data show opportunities for competitors to gain an advantage?



Audit of public geospatial information

RAND's 2004 deliverable included an audit of

- 465 programs/offices/initiatives at 30 agencies and departments identified as providing geospatial information to the public
  - 628 public datasets sampled from websites
  - 37 (~6%) found to be useful in helping an attacker select a target or plan an attack against a site
    - None were considered so critical that an "attacker could not perform the attack without" them
- Conclusions
  - Publically available geospatial "information needed for identifying and locating potential targets is widely accessible"
  - "…detailed and up-to-date information required for attack planning against a particular target is much less readily available"


Mapping the Risks

Assessing the Homeland Security Implications of Publicly Available Geospatial Information

JOHN C. BAKER, BETH E. LACHMAN, DAVID R. FRELINGER, KEVIN M. O'CONNELL, ALEXANDER C. HOU, MICHAEL S. TSENG, DAVID ORLETSKY, CHARLES YOST

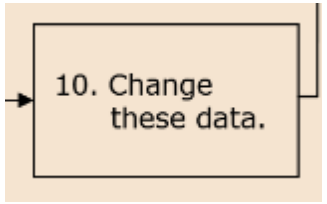
Prepared for the National Geospatial-Intelligence Agency

Approved for public release, distribution unlimited

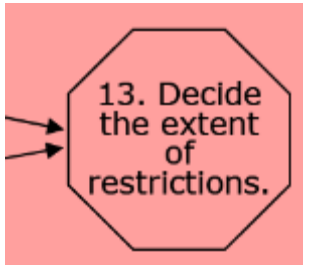
RAND NATIONAL DEFENSE RESEARCH INSTITUTE

If security risks outweigh benefits of releasing the data to the public, agency can choose to safeguard data by:

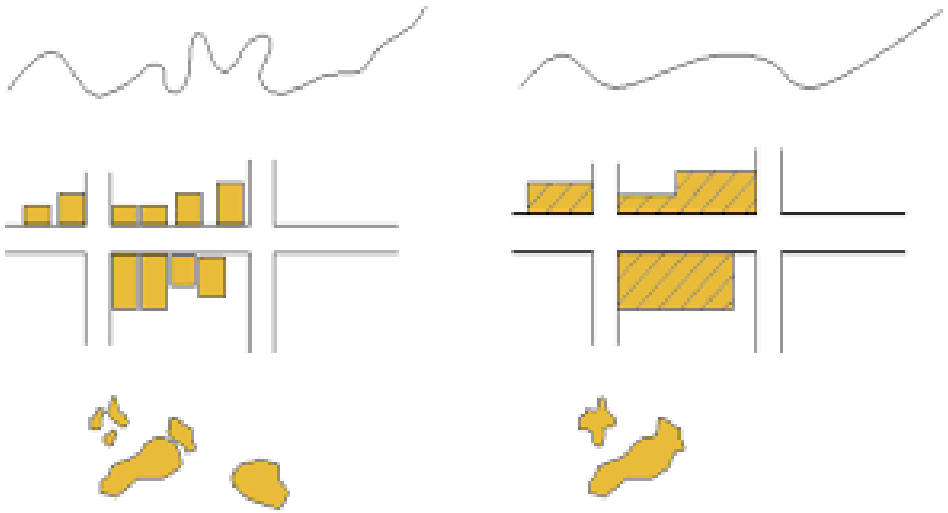
- **Modifying data**
  - Remove or reduce detail in offending data elements
    - either in the attributes, spatial representations, or both



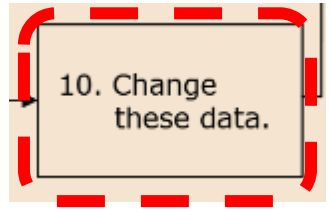
- **Restricting access to data**
  - If agency lacks authority to change data, or believes modifying data will undermine its value to the public, then agency can restrict access



…control/mitigate risk…



Before… *…after*



To remove or reduce detail in offending data elements apply techniques of ***Cartographic Generalization***

1. *Selective Omission*
2. *Simplification*
3. *Combination*

What is the security objective of FGDC's Guidelines ?

Confidentiality
Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information.

?

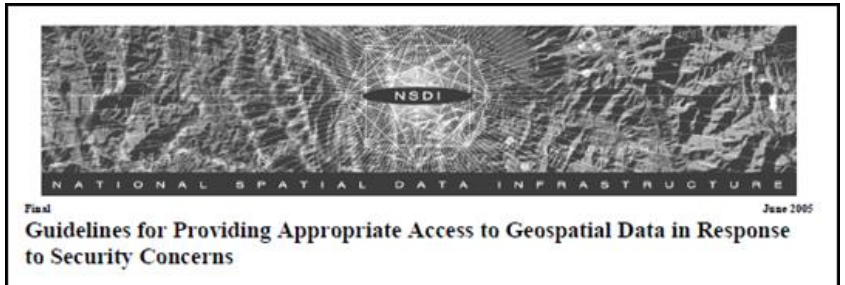
Integrity
Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity.

?

Availability
Ensuring timely and reliable access to and use of information.

?

What FIPS 199 security objectives are at risk by implementing the FGDC's Guidelines ?



Metadata enables communicating data classification information

2 examples of metadata standards that include security categorization information for geographic datasets

FGDC-STD-001-1998

National Spatial Data Infrastructure

Content Standard for Digital Geospatial Metadata

Metadata Ad Hoc Working Group
Federal Geographic Data Committee

Federal Geographic Data Committee
Department of Agriculture ● Department of Commerce ● Department of Defense ● Department of Energy
Department of Housing and Urban Development ● Department of the Interior ● Department of State
Department of Transportation ● Environmental Protection Agency
Federal Emergency Management Agency ● Library of Congress
National Aeronautics and Space Administration ● National Archives and Records Administration
Tennessee Valley Authority

EUROPEAN STANDARD
NORME EUROPÉENNE
EUROPÄISCHE NORM

EN ISO 19115-1

April 2014

ICS 35.240.70

Supersedes EN ISO 19115:2005

English Version

Geographic information — Metadata — Part 1: Fundamentals (ISO 19115-1:2014)

Information géographique — Métadonnées — Partie 1: Principes de base (ISO 19115-1:2014)

Geoinformation — Metadaten — Teil 1: Grundsätze (ISO 19115-1:2014)

This European Standard was approved by CEN on 22 February 2014.

CEN members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration. Up-to-date lists and bibliographical references concerning such national standards may be obtained on application to the CEN-CENELEC Management Centre or to any CEN member.

This European Standard exists in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CEN member into its own language and notified to the CEN-CENELEC Management Centre has the same status as the official versions.

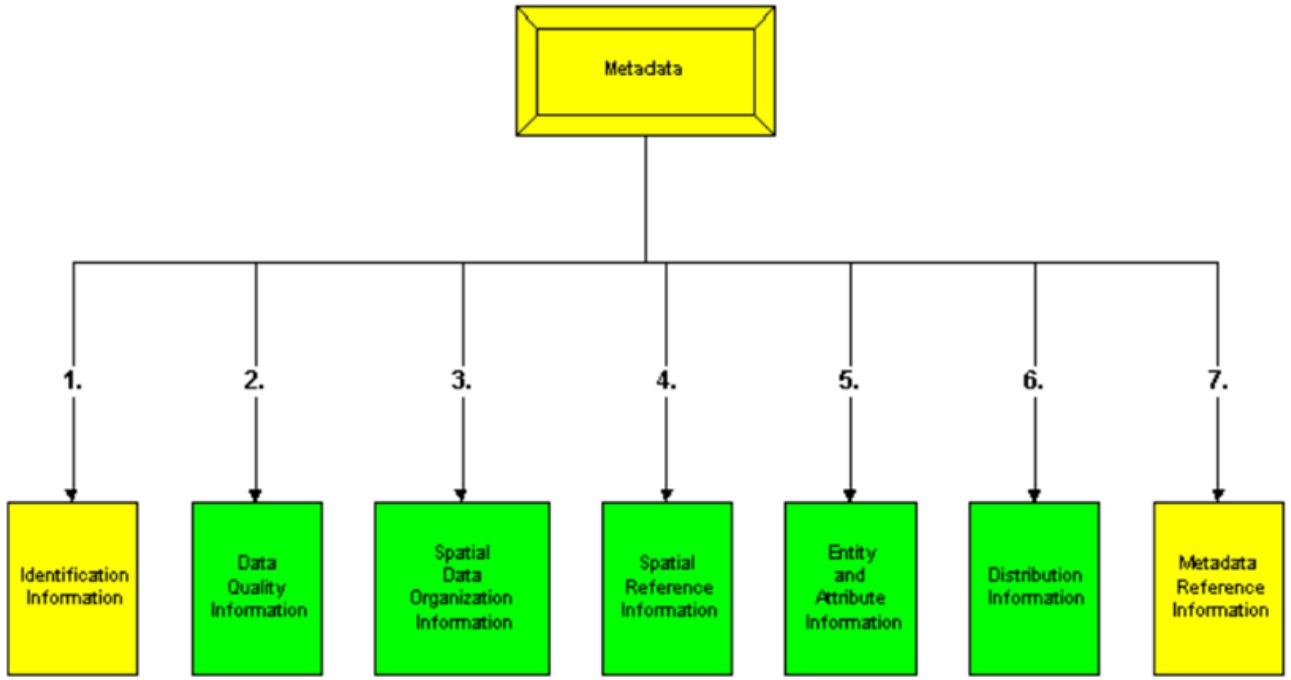
CEN members are the national standards bodies of Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Former Yugoslav Republic of Macedonia, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and United Kingdom.

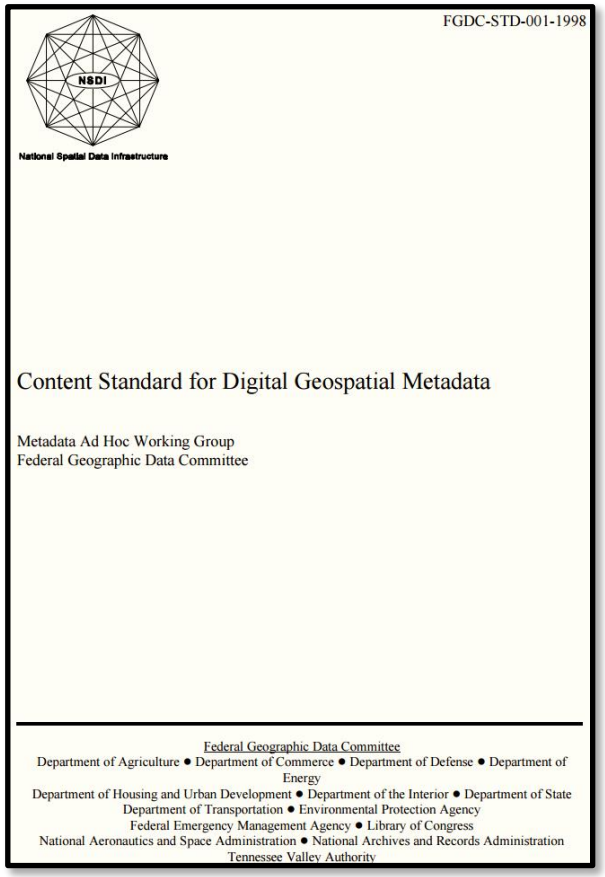
cen

EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

CEN-CENELEC Management Centre: Avenue Marnix 17, B-1000 Brussels

Ref. No. EN ISO 19115-1:2014 E





FGDC-STD-001-1998

NSDI

National Spatial Data Infrastructure

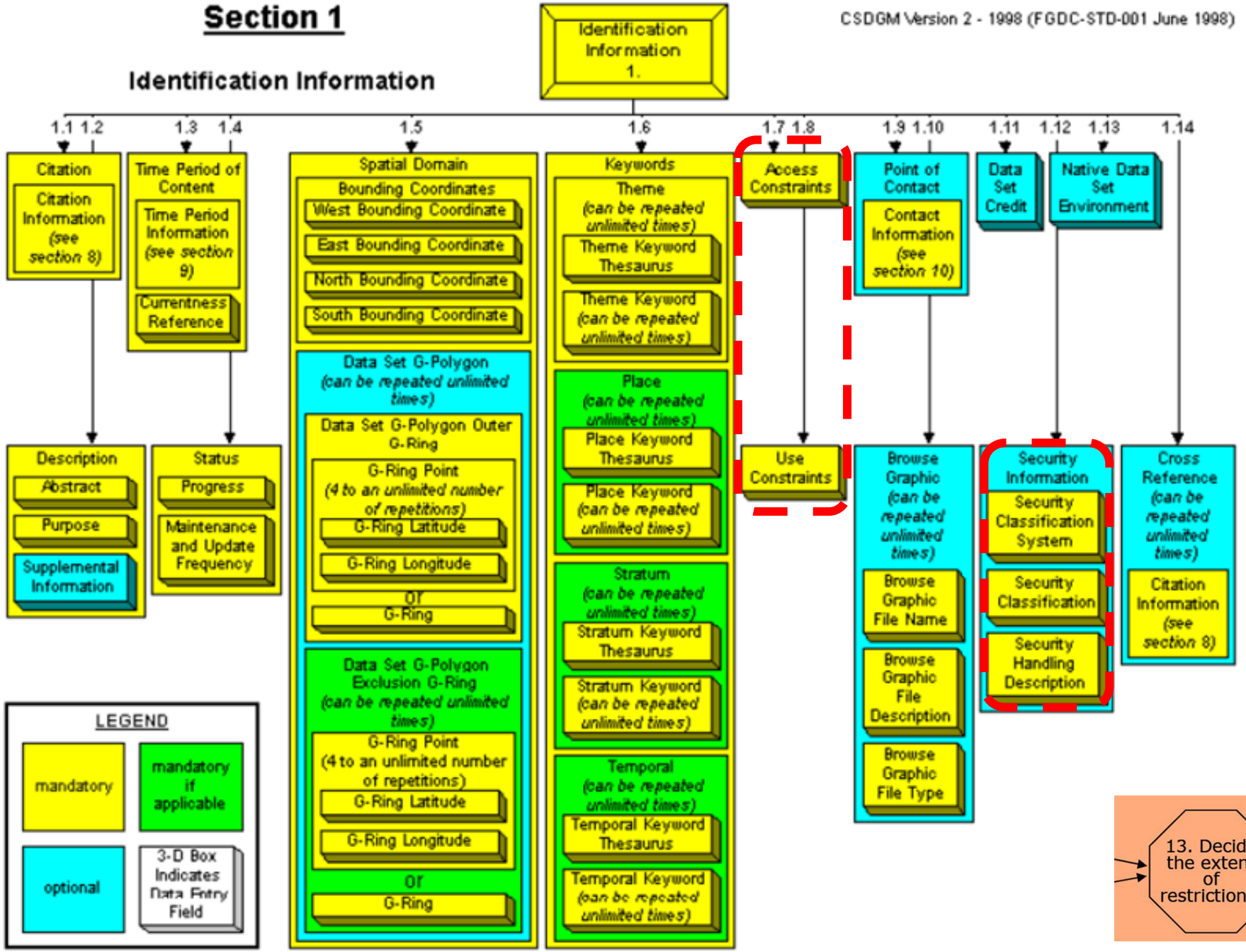
Content Standard for Digital Geospatial Metadata

Metadata Ad Hoc Working Group
Federal Geographic Data Committee

Federal Geographic Data Committee
Department of Agriculture ● Department of Commerce ● Department of Defense ● Department of Energy
Department of Housing and Urban Development ● Department of the Interior ● Department of State
Department of Transportation ● Environmental Protection Agency
Federal Emergency Management Agency ● Library of Congress
National Aeronautics and Space Administration ● National Archives and Records Administration
Tennessee Valley Authority

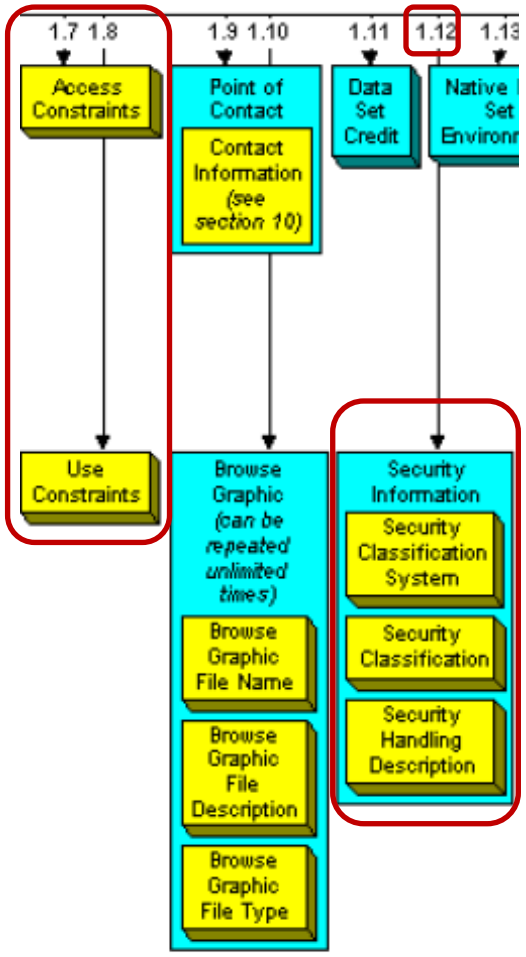
Section 1

Identification Information



13. Decide the extent of restrictions.

Communicating risk classification and controls…



Note: Be wary of *metadata with undefined or free text domains*

1.7 Access Constraints -- restrictions and legal prerequisites for accessing the data set. These include any access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the data set.
Type: text
Domain: "None" free text ←
Short Name: accconst

1.8 Use Constraints -- restrictions and legal prerequisites for using the data set after access is granted. These include any use constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on using the data set.
Type: text
Domain: "None" free text ←
Short Name: useconst

1.12 Security Information -- handling restrictions imposed on the data set because of national security, privacy, or other concerns.
Type: compound
Short Name: secinfo

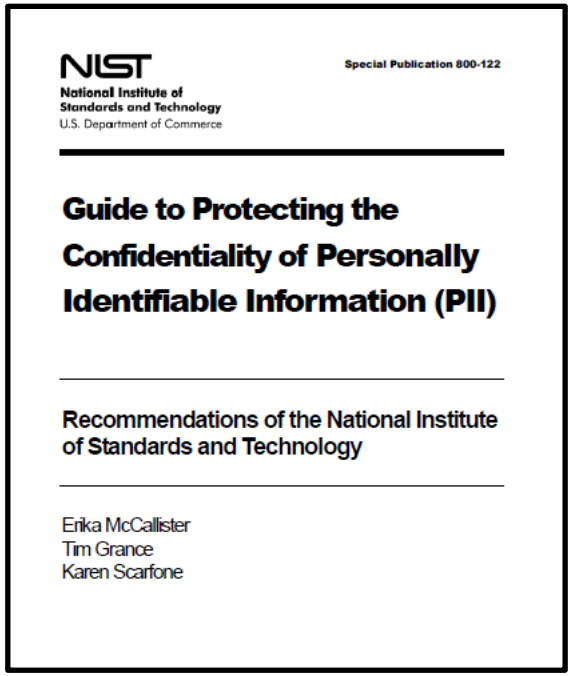
1.12.1 Security Classification System -- name of the classification system.
Type: text
Domain: free text ←
Short Name: secsys

1.12.2 Security Classification -- name of the handling restrictions on the data set.
Type: text
Domain: "Top secret" "Secret" "Confidential" "Restricted" "Unclassified" "Sensitive" ←
free text ←
Short Name: secclass

1.12.3 Security Handling Description -- additional information about the restrictions on handling the data set.
Type: text
Domain: free text ←
Short Name: sechandl

NIST SP 800-122 – Guide to Protecting Confidentiality of PII

- Specifically focused on:
  - Identifying PII
  - **Determining PII confidentiality** impact level needed to supplement the FIPS 199 confidentiality impact level of an information system

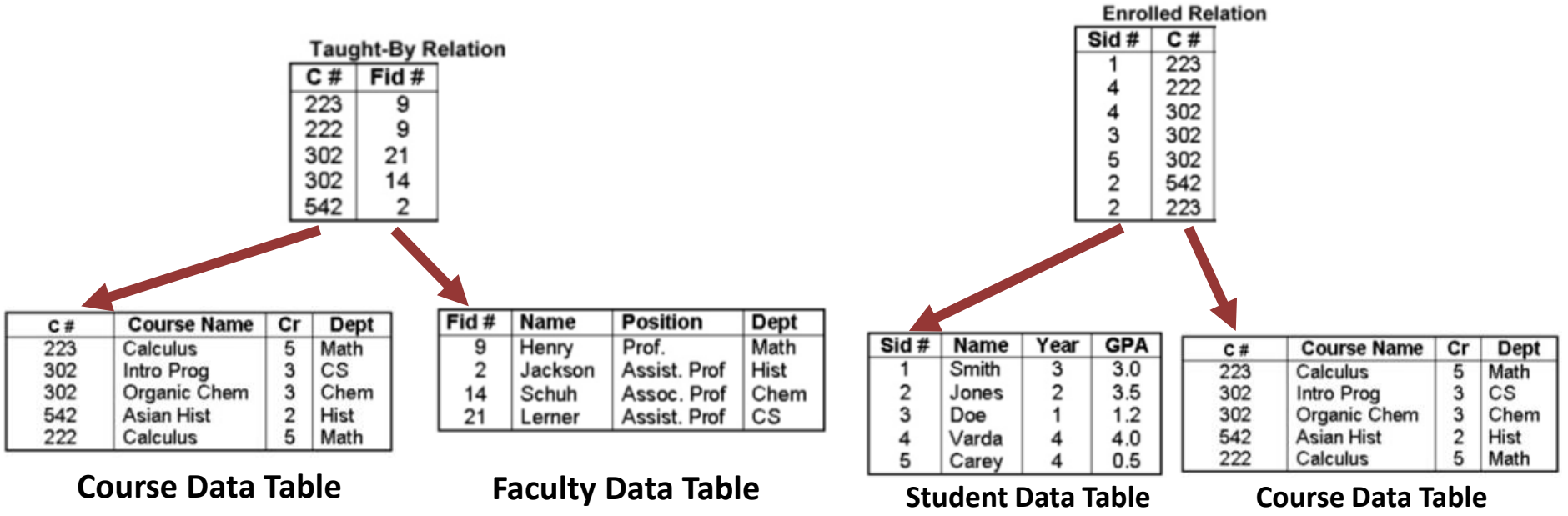


Personally Identifiable Information (PII)

Any information about an individual maintained by an agency, including:

1. Any information that can be used to distinguish (i.e. identify) or trace an individual's identity, such as:
   - *Name*
   - *Identifying number*
   - *Address*
   - *Asset identifier*
   - *Telephone number*
   - *Personal characteristics*
   - *Personally owned property identifiers*
2. Any other information that is linked or linkable to the identifiers listed in #1:
   - Date of birth
   - Place of birth
   - Race
   - Religion
   - Weight
   - Geographic indicators
   - Medical information
   - Educational information
   - Financial information
   - Employment information
   - …

Linked information



Taught-By Relation

C #	Fid #
223	9
222	9
302	21
302	14
542	2

C #	Course Name	Cr	Dept
223	Calculus	5	Math
302	Intro Prog	3	CS
302	Organic Chem	3	Chem
542	Asian Hist	2	Hist
222	Calculus	5	Math

Course Data Table

Fid #	Name	Position	Dept
9	Henry	Prof.	Math
2	Jackson	Assist. Prof	Hist
14	Schuh	Assoc. Prof	Chem
21	Lerner	Assist. Prof	CS

Faculty Data Table

Enrolled Relation

Sid #	C #
1	223
4	222
4	302
3	302
5	302
2	542
2	223

Sid #	Name	Year	GPA
1	Smith	3	3.0
2	Jones	2	3.5
3	Doe	1	1.2
4	Varda	4	4.0
5	Carey	4	0.5

Student Data Table

C #	Course Name	Cr	Dept
223	Calculus	5	Math
302	Intro Prog	3	CS
302	Organic Chem	3	Chem
542	Asian Hist	2	Hist
222	Calculus	5	Math

Course Data Table

Linkable information

Property ("Parcel") Data Table

Shape	ID	PIN	Area	Addr	Code
	1	334-1626-001	7,342	341 Cherry Ct.	SFR
	2	334-1626-002	8,020	343 Cherry Ct.	UND
	3	334-1626-003	10,031	345 Cherry Ct.	SFR
	4	334-1626-004	9,254	347 Cherry Ct.	SFR
	5	334-1626-005	8,856	348 Cherry Ct.	UND
	6	334-1626-006	9,975	346 Cherry Ct.	SFR
	7	334-1626-007	8,230	344 Cherry Ct.	SFR
	8	334-1626-008	8,645	342 Cherry Ct.	SFR

Owner Tax Data Table

PIN	Owner	Acq.Date	Assessed	TaxStat
334-1626-001	G. Hall	1995/10/20	$115,500.00	02
334-1626-002	H. L Holmes	1993/10/06	$24,375.00	01
334-1626-003	W. Rodgers	1980/09/24	$175,500.00	02
334-1626-004	J. Williamson	1974/09/20	$135,750.00	02
334-1626-005	P. Goodman	1966/06/06	$30,350.00	02
334-1626-006	K. Staley	1942/10/24	$120,750.00	02
334-1626-007	J. Dormandy	1996/01/27	$110,650.00	01
334-1626-008	S. Gooley	2000/05/31	$145,750.00	02

PIN is a common identifying number that can serve as a "foreign key" to link the data tables together

Personally Identifiable Information (PII)

Any information about an individual maintained by an agency, including:

1. Any information that can be used to distinguish (i.e. identify) or trace an individual's identity, such as:
   - *Name*
   - *Identifying number*
   - *Address*
   - *Asset identifier*
   - *Telephone number*
   - *Personal characteristics*
   - *Personally owned property identifiers*

2. Any other information that is linked or linkable to the identifiers listed in #1:
   - Date of birth
   - Place of birth
   - Race
   - Religion
   - Weight
   - Geographic indicators
   - Medical information
   - Educational information
   - Financial information
   - Employment information
   - …

Is this PII ?

Property ("Parcel") Data Table

Shape	ID	PIN	Area	Addr	Code
	1	334-1626-001	7,342	341 Cherry Ct.	SFR
	2	334-1626-002	8,020	343 Cherry Ct.	UND
	3	334-1626-003	10,031	345 Cherry Ct.	SFR
	4	334-1626-004	9,254	347 Cherry Ct.	SFR
	5	334-1626-005	8,856	348 Cherry Ct.	UND
	6	334-1626-006	9,975	346 Cherry Ct.	SFR
	7	334-1626-007	8,230	344 Cherry Ct.	SFR
	8	334-1626-008	8,645	342 Cherry Ct.	SFR

Owner Tax Data Table

PIN	Owner	Acq.Date	Assessed	TaxStat
334-1626-001	G. Hall	1995/10/20	$115,500.00	02
334-1626-002	H. L Holmes	1993/10/06	$24,375.00	01
334-1626-003	W. Rodgers	1980/09/24	$175,500.00	02
334-1626-004	J. Williamson	1974/09/20	$135,750.00	02
334-1626-005	P. Goodman	1966/06/06	$30,350.00	02
334-1626-006	K. Staley	1942/10/24	$120,750.00	02
334-1626-007	J. Dormandy	1996/01/27	$110,650.00	01
334-1626-008	S. Gooley	2000/05/31	$145,750.00	02

Agenda

- ✓Vocabulary
- ✓Data Classification Process and Models
- Test taking tip
- Quiz

Test Taking Tip

- Read the answers first -

This contradicts many people's test taking recommendations…

…but, it works. Here's why:

- Quickly alerts you to the type of question to expect
- Focuses your attention in reading the question for meaningful information
- Gives you advanced warning that there may be more than one significant concepts (option to answer in the form "Both A & B")
- Gives you an opportunity to get a sense of the sort of answer the test maker is looking for
- There may be more than one valid answer, but the test maker may be looking for "best mitigation for the situation" or "least risk in the situation"

Test Taking Tip

- Read the answers first -

This contradicts many people's test taking recommendations…

…but, it works. Here's why:

- Quickly alerts you to the type of question to expect
- Focuses your attention in reading the question for meaningful information
- Gives you advanced warning that there may be more than one significant concepts (option to answer in the form "Both A & B")
- Gives you an opportunity to get a sense of the sort of answer the test maker is looking for
- There may be more than one valid answer, but the test maker may be looking for "best mitigation for the situation" or "least risk in the situation"

Test Taking Tip

Example:

A. Transaction authorization
B. Loss or duplication of EDI transmissions
C. Transmission delay
D. Deletion or manipulation of transactions prior to or after establishment of application controls

Test Taking Tip

Example:

Which of the following represents the GREATEST potential risk in an Electronic Data Interchange (EDI) environment?

A. Transaction authorization
B. Loss or duplication of EDI transmissions
C. Transmission delay
D. Deletion or manipulation of transactions prior to or after establishment of application controls

Test Taking Tip

Example:

Which of the following represents the GREATEST potential risk in an Electronic Data Interchange (EDI) environment?

A. Transaction authorization
B. Loss or duplication of EDI transmissions
C. Transmission delay
D. Deletion or manipulation of transactions prior to or after establishment of application controls

Answer: A

Quiz

1. Which of the choices below is the most often used criteria to determine the classification of a business object?

   a. Value
   b. Useful life
   c. Age
   d. Personal association

Quiz

1. Which of the choices below is the most often used criteria to determine the classification of a business object?

   a. Value
   b. Useful life
   c. Age
   d. Personal association

Quiz

2. Which of the below definitions is the best description of a vulnerability?

a. A weakness in a system that could be exploited
b. A company resource that is lost due to an incident
c. The minimum loss associated with an incident
d. A potential incident that could cause harm

Quiz

2. Which of the below definitions is the best description of a vulnerability?

a. A weakness in a system that could be exploited
b. A company resource that is lost due to an incident
c. The minimum loss associated with an incident
d. A potential incident that could cause harm

Quiz

5. Which group represents the most likely source of an asset loss through in appropriate computer use?

a. Crackers
b. Hackers
c. Employees
d. Saboteurs

Quiz

5. Which group represents the most likely source of an asset loss through in appropriate computer use?

a. Crackers
b. Hackers
c. Employees
d. Saboteurs

Agenda

- ✓Vocabulary
- ✓Data Classification Process and Models
- ✓Test taking tip
- ✓Quiz