

Guidelines for Data Classification

Document Information	
Status	Published
Published	09/15/2009
Last Updated	09/15/2011
Version	1.0

Revision History

Version	Published	Author	Description
0.1	07/02/2008	Doug Markiewicz	Original draft
0.2	09/25/2008	Doug Markiewicz	Replaced Categorization section with Data Collections and added sections on Reclassification and Calculating Classifications
0.3	10/20/2008	Doug Markiewicz	Rewrote section on Calculating Classifications due to flaws in original system. Updated Purpose, Applies To and Definitions.
0.4	11/04/2008	Doug Markiewicz	Removed equation, made a minor update to the definition of Public Data and updated Additional Information. Sorted Appendix A so that terms appear in alphabetical order and added Covered Financial Information as a term.
0.5	02/20/2009	Doug Markiewicz	Added a missing bullet to the last identifier in listed in Appendix A definition G. The definition itself was not modified.
0.6	02/26/2009	Doug Markiewicz	Various updates based on feedback provided by Mary Ann Blair. Major changes include, adding 'Data Steward' to the Definitions, adding references to Information Security Roles & Responsibilities and adding Federal Tax Information to Appendix A.
0.7	03/18/2009	Doug Markiewicz	Updated definition of PHI in Appendix A to reference the HIPAA Information Security Policy. Added Authentication Verifier to Appendix A.
0.8	09/15/2009	Doug Markiewicz	Updated Applied To for consistency with related publications. Removed Education Records from Appendix A per the recommendation of General Counsel. Updated Personally Identifiable Education Records in Appendix A to reference the Policy on Student Privacy Rights.
0.9	01/22/2010	Doug Markiewicz	Updated Appendix A to include Export Controlled Materials.
1.0	09/15/2011	Doug Markiewicz	Updated definition of Protected Health Information to align with the new HIPAA Policy. Removed DRAFT designation.



Purpose

The purpose of this Guideline is to establish a framework for classifying institutional data based on its level of sensitivity, value and criticality to the University as required by the University's Information Security Policy. Classification of data will aid in determining baseline security controls for the protection of data.

Applies To

This Policy applies to all faculty, staff and third-party Agents of the University as well as any other University affiliate who is authorized to access Institutional Data. In particular, this Guideline applies to those who are responsible for classifying and protecting Institutional Data, as defined by the [Information Security Roles and Responsibilities](#).

Definitions

Confidential Data is a generalized term that typically represents data classified as Restricted, according to the data classification scheme defined in this Guideline. This term is often used interchangeably with *sensitive data*.

A *Data Steward* is a senior-level employee of the University who oversees the lifecycle of one or more sets of Institutional Data. See the [Information Security Roles and Responsibilities](#) for more information.

Institutional Data is defined as all data owned or licensed by the University.

Non-public Information is defined as any information that is classified as Private or Restricted Information according to the data classification scheme defined in this Guideline.

Sensitive Data is a generalized term that typically represents data classified as Restricted, according to the data classification scheme defined in this Guideline. This term is often used interchangeably with *confidential data*.

Data Classification

Data classification, in the context of information security, is the classification of data based on its level of sensitivity and the impact to the University should that data be disclosed, altered or destroyed without authorization. The classification of data helps determine what baseline security controls are appropriate for safeguarding that data. All institutional data should be classified into one of three sensitivity levels, or classifications:

A. Restricted Data

Data should be classified as Restricted when the unauthorized disclosure, alteration or destruction of that data could cause a significant level of risk to the University or its affiliates. Examples of Restricted data include data protected by state or federal privacy regulations and data protected by confidentiality agreements. The highest level of security controls should be applied to Restricted data.

B. Private Data

Data should be classified as Private when the unauthorized disclosure, alteration or destruction of that data could result in a moderate level of risk to the University or its affiliates. By default, all Institutional Data that is not explicitly classified as Restricted or Public data should be treated as Private data. A reasonable level of security controls should be applied to Private data.

C. Public Data

Data should be classified as Public when the unauthorized disclosure, alteration or destruction of that data would result in little or no risk to the University and its affiliates. Examples of Public data include press releases, course information and research publications. While little or no controls are required to protect the confidentiality of Public data, some level of control is required to prevent unauthorized modification or destruction of Public data.

Classification of data should be performed by an appropriate Data Steward. Data Stewards are senior-level employees of the University who oversee the lifecycle of one or more sets of Institutional Data. See [Information Security Roles and Responsibilities](#) for more information on the Data Steward role and associated responsibilities.

Data Collections

Data Stewards may wish to assign a single classification to a collection of data that is common in purpose or function. When classifying a collection of data, the most restrictive classification of any of the individual data elements should be used. For example, if a data collection consists of a student's name, address and social security number, the data collection should be classified as Restricted even though the student's name and address may be considered Public information.

Reclassification

On a periodic basis, it is important to reevaluate the classification of Institutional Data to ensure the assigned classification is still appropriate based on changes to legal and contractual obligations as well as changes in the use of the data or its value to the University. This evaluation should be conducted by the appropriate Data Steward. Conducting an evaluation on an annual basis is encouraged; however, the Data Steward should determine what frequency is most appropriate based on available resources. If a Data Steward determines that the classification of a certain data set has changed, an analysis of security controls should be performed to determine whether existing controls are consistent with the new classification. If gaps are found in existing security controls, they should be corrected in a timely manner, commensurate with the level of risk presented by the gaps.

Calculating Classification

The goal of information security, as stated in the University's Information Security Policy, is to protect the confidentiality, integrity and availability of Institutional Data. Data classification reflects the level of impact to the University if confidentiality, integrity or availability is compromised.

Unfortunately there is no perfect quantitative system for calculating the classification of a particular data element. In some situations, the appropriate classification may be more obvious, such as when federal laws require the University to protect certain types of data (e.g. personally identifiable information). If the appropriate classification is not inherently obvious, consider each security objective using the following table as a guide. It is an excerpt from [Federal Information Processing Standards \("FIPS"\) publication 199](#) published by the National Institute of Standards and Technology, which discusses the categorization of 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.	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.	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.	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.

As the total potential impact to the University increases from Low to High, the classification of data should become more restrictive moving from Public to Restricted. If an appropriate classification is still unclear after considering these points, contact the Information Security Office for assistance.

Additional Information

If you have any questions or comments related to these Standards, please send email to the University Information Security Office at iso@andrew.cmu.edu.

Additional information can also be found using the following resources:

- Information Security Policy
<http://www.cmu.edu/iso/governance/policies/information-security.html>
- Information Security Roles and Responsibilities
<http://www.cmu.edu/iso/governance/policies/information-security-roles.html>
- Policy on Student Privacy Rights
<http://www.cmu.edu/policies/documents/StPrivacy.html>
- Gramm-Leach-Bliley Information Security Program
<http://www.cmu.edu/policies/documents/ISP.htm>
- HIPAA Policy
<http://www.cmu.edu/policies/documents/HIPAA.htm>
- FIPS Publication 199: Standards for Security Categorization
<http://csrc.nist.gov/publications/fips/fips199/FIPS-PUB-199-final.pdf>
- Internal Revenue Service Publication 1075: Tax Information Security Guidelines
<http://www.irs.gov/pub/irs-pdf/p1075.pdf>

Appendix A – Predefined Types of Restricted Information

The Information Security Office and the Office of General Counsel have defined several types of Restricted data based on state and federal regulatory requirements. They're defined as follows:

1. Authentication Verifiers

An Authentication Verifier is a piece of information that is held in confidence by an individual and used to prove that the person is who they say they are. In some instances, an Authentication Verifier may be shared amongst a small group of individuals. An Authentication Verifier may also be used to prove the identity of a system or service. Examples include, but are not limited to:

- Passwords
- Shared secrets
- Cryptographic private keys

2. Covered Financial Information

See the University's [Gramm-Leach-Bliley Information Security Program](#).

3. Electronic Protected Health Information ("EPHI")

EPHI is defined as any Protected Health Information ("PHI") that is stored in or transmitted by electronic media. For the purpose of this definition, electronic media includes:

- Electronic storage media includes computer hard drives and any removable and/or transportable digital memory medium, such as magnetic tape or disk, optical disk, or digital memory card.
- Transmission media used to exchange information already in electronic storage media. Transmission media includes, for example, the Internet, an extranet (using Internet technology to link a business with information accessible only to collaborating parties), leased lines, dial-up lines, private networks and the physical movement of removable and/or transportable electronic storage media. Certain transmissions, including of paper, via facsimile, and of voice, via telephone, are not considered to be transmissions via electronic media because the information being exchanged did not exist in electronic form before the transmission.

4. Export Controlled Materials

Export Controlled Materials is defined as any information or materials that are subject to United States export control regulations including, but not limited to, the Export Administration Regulations ("EAR") published by the U.S. Department of Commerce and the International Traffic in Arms Regulations ("ITAR") published by the U.S. Department of State. See the [Information and Guidelines on Federal Export Control Laws and Regulations](#), published by the [Office of Sponsored Programs](#), for more information.

5. Federal Tax Information ("FTI")

FTI is defined as any *return, return information or taxpayer return information* that is entrusted to the University by the Internal Revenue Services. See [Internal Revenue Service Publication 1075 Exhibit 2](#) for more information.

6. Payment Card Information

Payment card information is defined as a credit card number (also referred to as a primary account number or PAN) in combination with one or more of the following data elements:

- Cardholder name
- Service code
- Expiration date
- CVC2, CVV2 or CID value
- PIN or PIN block
- Contents of a credit card's magnetic stripe

7. Personally Identifiable Education Records

Personally Identifiable Education Records are defined as any Education Records that contain one or more of the following personal identifiers:

- Name of the student
- Name of the student's parent(s) or other family member(s)
- Social security number
- Student number
- A list of personal characteristics that would make the student's identity easily traceable
- Any other information or identifier that would make the student's identity easily traceable

See Carnegie Mellon's [Policy on Student Privacy Rights](#) for more information on what constitutes an Education Record.

8. Personally Identifiable Information ("PII")

For the purpose of meeting state security breach notification requirements, PII is defined as a person's first name or first initial and last name in combination with one or more of the following data elements:

- Social security number
- State-issued driver's license number
- State-issued identification card number
- Financial account number in combination with a security code, access code or password that would permit access to the account
- Medical and/or health insurance information

9. Protected Health Information ("PHI")

PHI is defined as "individually identifiable health information" transmitted by electronic media, maintained in electronic media or transmitted or maintained in any other form or medium by a Covered Component, as defined in Carnegie Mellon's [HIPAA Policy](#). PHI is considered individually identifiable if it contains one or more of the following identifiers:

- Name
- Address (all geographic subdivisions smaller than state including street address, city, county, precinct or zip code)
- All elements of dates (except year) related to an individual including birth date, admissions date, discharge date, date of death and exact age (if over 89)
- Telephone numbers

Carnegie Mellon

- Fax numbers
- Electronic mail addresses
- Social security numbers
- Medical record numbers
- Health plan beneficiary numbers
- Account numbers
- Certificate/license numbers
- Vehicle identifiers and serial numbers, including license plate number
- Device identifiers and serial numbers
- Universal Resource Locators (URLs)
- Internet protocol (IP) addresses
- Biometric identifiers, including finger and voice prints
- Full face photographic images and any comparable images
- Any other unique identifying number, characteristic or code that could identify an individual

Per Carnegie Mellon's [HIPAA Policy](#), PHI does not include education records or treatment records covered by the Family Educational Rights and Privacy Act or employment records held by the University in its role as an employer.