Manage risk with the classification and protection of digital research data

Trusted CI Fellows Report

March 16, 2022

For public release

Author, Michael Kyle
Trusted CI Fellow
University of Delaware

Data is a necessity for research. Without data to analyze, researchers would not be able to prove hypotheses, innovate new ideas, and have their results validated by their peers. Therefore, data security is a critical part of research. While not all research data contains confidential or proprietary information, all research does require data security to maintain its integrity. When data is compromised, either accidentally or from malicious actors, the effects can be serious. In extreme cases there could be legal consequences for the leaking of improperly stored data. In all cases, compromised research integrity leads to lost work, time, and money.

This paper is a guide for researchers at the University of Delaware (UD) and their collaborators and partners and will be used to summarize the materials and tools provided by Secure UD. The aim is to assist researchers with classifying their data appropriately, and to guide them in using best practices to protect their data. This paper will specifically focus on digital data. Though some of the material covered might be applicable to non-digital data, that is not the primary focus.

Research data is defined as the "recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer-reviews, or communications with colleagues (OMB Circular 110). For digital data, this can include but is not limited to digital documents, survey results, slides, or questionnaires. It can also include databases, images, videos and audio files, instrumentation input and output files, models, algorithms, scripts, or proprietary software. Research data does not include trade secrets, information considered confidential like personnel and Protected Health Information (PHI), or other information that is protected under federal and local privacy laws.

Research data comes in many different formats and sizes. It also comes with various levels of sensitivity. At UD, there are three levels of data and information classifications. They are outlined in Table 1. This table is based on Secure UD's University Information Classifications. The examples have been slightly modified to reflect common types of research data.

To properly protect data, researchers need to first identify the classification level of the data. Once the risk level of data has been correctly classified, actions to minimize the risk can begin. The simplest way to start this process is to use the Secure UD Research Security Plan Tool and Secure UD Managing Research Data Risks questionnaire. These two documents ask a series of

questions about the research environment and the data used in the research. They are designed to help researchers identify risks associated with the research and then develop controls to mitigate those risks. Once the two documents are completed, it provides researchers with tools and a plan for securing research data.

The Secure UD Research Security Plan Tool starts off by asking researchers to briefly describe their research and the classification level of the data. From there the tool asks “yes” or “no” questions about Physical Risk, Confidential Risk, Integrity Risk, Availability Risk, Privacy Risk, Human Risk and Legal, Regulatory and Contractual Risk. Researchers’ responses guide them in identifying controls to manage risks associated with their data. It also helps in documenting and developing a security plan around those risks. The security plan will be unique to the risks the researchers identified for their research. The plan can then be paired with the Secure UD Managing Research Data Risks questionnaire.

The UD Managing Research Data Risks questionnaire can be used for taking a second look at the risks the research data is facing and help in the development of controls for the research security plan. Some of the questions may be slightly redundant to those asked in the Secure UD Research Security Plan Tool, but they are more generic. The questionnaire can be answered by anyone associated with the research project. However, the security plan tool is geared towards the Principal Investigator (PI) of the research project answering the questions. The benefit of using the questionnaire is that each of the different sections of risk provide steps that researchers can use to help mitigate that particular risk. If all members associated with the research project answer the questionnaire, it also helps the PI to identify areas of risk not previously defined.

In summary, Secure UD provides tools to help researchers identify and manage risks associated with research and its data. In addition to these tools, researchers may ask for additional consultation with IT’s Secure UD and Research Cyberinfrastructure teams. The world of cybersecurity is constantly evolving, and new threats and vulnerabilities are always being developed and exploited. Therefore, it is important for researchers to constantly evaluate their controls and security plans to adapt to new threats as they come.

Table 1
Research Data Classification

Level I Low Risk	Level II Moderate Risk	Level III High Risk
Risk: Unintentional, unlawful, or unauthorized disclosure presents **limited or no risk**.	**Risk:** Unintentional, unlawful, or unauthorized disclosure presents **moderate risk.**	**Risk:** Unintentional, unlawful, or unauthorized disclosure presents **significant risk**.
Examples: • Data gathered from the public domain • Data with low confidentiality, integrity, or availability concerns • Data not under contractual, legal, or other requirements	**Examples:** • FERPA records • Unpublished intellectual property • Other data specified by contractual, legal, or other requirements	**Examples:** • Personally Identifiable Information (PII) • Protected health information (PHI/ePHI) • Export-restricted data • Human subject data • UDelNet passwords • Encryption keys • Other data specified by contractual, legal, or other requirements
Protection requirements: • May be shared publicly	**Protection requirements:** • Share only with those who need to know	**Protection requirements:** • Encrypt at rest and in transit • Access, process, store, and transmit only using managed computers • Do not use cloud services to access, process, store, or transmit unless those services are explicitly approved for that data

Source: University Information Classifications

a. This table is based on Secure UD's *University Information Classifications.* The examples have been slightly modified to reflect common types of research data. To properly protect data, first identify the classification level of the data.

Works Cited

Chapter II—Office of Management and Budget Circulars and ...
https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A110/2cfr215-0.pdf.
Secure UD Research Security Plan Tool, University of Delaware Information Technologies IT Security, 2020, https://www1.udel.edu/security/resources/downloads/Secure%20UD%20Research%20Security%20Plan%20Tool.docx.
"UD Managing Research Data Risks." *Secure UD Managing Research Data Risks.pdf*, University of Delaware Information Technologies IT Security, 2020, https://www1.udel.edu/security/resources/downloads/Secure%20UD%20Managing%20Research%20Data%20Risks.pdf.
University of Delaware Information Technologies IT Security. *University Information Classifications*, 2020, https://www1.udel.edu/security/framework/classification.html.
"Welcome to Secure UD." *Secure UD Home*, https://www1.udel.edu/security/.