Bates College Data Management Guidelines

1. Purpose

These guidelines establish best practices for managing institutional data, distinguishing between data that should be stored on on-premises shared storage and data that can be stored in cloud-based solutions such as Google Drive.

2. Scope

These guidelines provide technical guidance on the long-term storage, organization, and management of institutional data at Bates College. They outline best practices for determining appropriate storage solutions, managing access, and ensuring the proper retention and disposal of data. This document is intended to support faculty, staff, and students who handle data throughout its lifecycle, ensuring data is securely stored, retained as necessary, and disposed of according to institutional policies and regulations.

3. Data Classification

College data is categorized into different sensitivity levels to ensure appropriate management and security measures are applied. These classifications consider legal requirements, contractual obligations, and ethical considerations, guiding how data should be protected to prevent unauthorized access. For more detailed information on the classification levels—Restricted, Internal, and Public—and specific guidelines for their use and security measures, please refer to the Data Classification Guidelines.

3.1. Classification Levels Overview:

3.1.1. **Restricted** - College data classified as Restricted include:

- Data protected specifically by federal or state law, or
- Data protected by college policy
- Data elements Data elements identified by the college as sensitive or confidential, even if not governed by external legal or regulatory requirements.

Additional Clarification on AI Usage:

- **Restricted data** (such as Personally Identifiable Information (PII), FERPA-protected information, health records, salary data, and financial aid information) **must not be processed or analyzed using AI tools** such as Gemini, ChatGPT, or similar platforms, due to the potential risks involved in AI data processing. Users must ensure that Restricted data are handled in compliance with the College's AI usage guidelines outlined in the Data Usage Policy.

3.1.2. **Internal** - College data are classified as Internal if they are not considered to be Restricted, and:

- The data are not generally available to the public, and
- The data are protected due to proprietary, ethical, or privacy considerations, even though there may not be a direct regulatory or common-law basis for requiring this protection.

3.1.3. **Public** - College data not otherwise identified as Internal or Restricted, and:

- The data is intended for public disclosure, or
- The loss of data would have no adverse impact on our mission, finances, or reputation.

4. Data Storage Locations

The college provides data storage for our end users through both on-campus shared drives and Google Drive (our primary cloud storage solution). Both storage options - Google Drive and On-Campus Shared Drives - are authorized to store all classification levels of data, including Restricted, Internal, and Public data, provided that the appropriate security measures, such as encryption and access controls, are applied.

4.1. On Campus Shared Drives

The college provides space for data storage on campus through various on-campus shared drives. These drives are securely backed up and encrypted at rest, ensuring the safety and integrity of the data. Access to these drives is restricted to individuals connected to the Bates network, and off-campus access requires the use of the Bates VPN.

- **Etna (Faculty / Academic Staff)**
  - Etna is designated for storing:
  - Scholarship data
  - Research data
  - Academic course data
- **China (Faculty / Academic Staff)**
  - China is designated for storing departmental data for faculty
- **Belfast (All Staff)**
  - Belfast is designated for storing:
  - Departmental data for staff
  - Files that are appropriate for retention
- **Paris**
  - Paris provides "home" or "personal" drives accessible to all Bates users.

4.2. Cloud Storage (e.g., Google Drive)

The college provides space for data storage in the cloud (Google Drive). Google Drive can be accessed from anywhere as long as the user has their Bates network credentials.

- **My Drive**

  should be used for:
  - Work related documents that you consider to be “yours” and would not be needed by co-workers should you leave the institution
  - Working documents related to your current projects and tasks.
  - Notes and individual work items that are part of your job responsibilities.

- **Shared Drive**

  should be used for:
  - Data actively being worked on or with a current business need
    - All classifications of data related to students, faculty, staff, alumnae, and College business can be stored in Google Shared Drive.
  - Collaborative project files
  - Files shared with individuals without Bates network credentials

4.3. Deciding Where to Store Institutional Data

When deciding where to store institutional data, consider the following factors:

- **Use Google Drive**: Use cloud storage (Google Drive) for collaborative projects, active working documents, or data that requires regular access from multiple locations. Google Drive is ideal for ongoing projects, enabling shared access and easy collaboration.
- **Use On-Premises Shared Drives**: Store data on Bates' on-premises shared drives for information that requires higher security or is not actively being used for collaboration. This includes sensitive data that needs to be protected by additional access controls or data that does not need to be accessed remotely. Shared drives also offer robust backup and disaster recovery features.
- **Security Considerations**: Both cloud and on-premises storage locations are authorized to store all data classification levels, but ensure that appropriate security measures (e.g., encryption, access controls) are applied based on the sensitivity of the data.

4.4. Managing Google Shared Drives

Bates community members can create and manage Shared Drives, controlling membership and permissions. When managing members, use the principle of least privilege by designating members as Content Managers, which provides them with limited control. Shared Drives can share information with both Bates and non-Bates accounts.

The “Shared with me” feature in Google Drive should be used sparingly, if at all. You lose access to a document that is shared with you when the owner of the document leaves Bates. This does not happen when that same file is in a Shared Drive.

4.5. Other Cloud Storage Providers

No institutional data (Restricted or Internal) may be stored on any cloud storage provider other than your Bates Google Shared Drive, with the following exception for faculty research data:
Faculty research data may be stored in other cloud storage providers, such as discipline-specific repositories. In such cases, the institution must be granted access to these data. This ensures that Bates College retains the ability to access institutional data as needed.Data Management Best Practices

4.6. For All Storage Locations

- Regularly review and remove unnecessary/redundant data.

4.7. For Cloud Storage (Google Drive)

- Use only institution-approved cloud services
- Regularly review sharing settings to ensure appropriate access
- Set expiration dates for shared links to limit access duration to sensitive documents
- Ensure that collaborative projects use Shared Drives instead of personal My Drives to maintain proper access control and data ownership
- Be cautious when sharing institutional data with individuals outside the college. Verify the necessity and appropriateness of external sharing

4.8. For Email Storage (Gmail)

- Regularly archive or delete unnecessary emails, especially those containing restricted information.
- Be mindful of forwarding emails containing restricted information.
- Before sending sensitive attachments, encrypt them using methods like password-protected Adobe PDFs or password protected Microsoft Office file formats (Word, Excel, Access, PowerPoint).
- Regularly review which apps and devices have access to your Gmail account and remove any that are no longer necessary.
- **Delete sensitive or restricted data** from your sent or received folders once the information has been transmitted and is no longer needed. This reduces the risk of leaving sensitive data stored in your email account indefinitely.

5. Data Transfer and Sharing

- Obtain appropriate approvals before sharing internal or restricted data

- Always encrypt restricted data before sending it to a third party or another institution. Best practice is to share the password for de-encryption via a phone call or a password manager.
- When data transmission is permitted, it must occur through secure channels to safeguard against unauthorized access or interception. For example:

  Use HTTPS: When uploading or downloading files from websites, make sure the website address starts with "https://".

  Use SFTP: For transferring files directly from one computer to another, use Secure File Transfer Protocol (SFTP), for example, CoreFTP.

  Remove sensitive data from your sent or received folder after confirming successful transmission, especially if it's restricted information.

6. Research Data Management

These guidelines work in conjunction with our institution's Research Data Management Policy, which can be found at:
https://www.bates.edu/ils/policies/access-use/research-data-management-policy-2/

All faculty, staff, and students engaged in research activities should familiarize themselves with both these guidelines and the Research Data Management Policy.

7. Training and Support

Contact the Help Desk for training and support with respect to data storage and management questions/needs.

8. Related Guidelines and Policies

These Data Management Guidelines should be read and applied in conjunction with the following related documents:

Research Data Management Policy
Acceptable Use Policy
Privacy Policy
Bates College Data Usage Policy (in progress Data Governance Committee)