March 20, 2024 | 10:30am-12pm
Virtual Workshop + Sandbox Session

How to Implement Encryption to Protect Your Research Data

u.mcmaster.ca/scds-events

RDM

SCDS

Library

McMaster University

How to Implement Encryption to Protect your Research Data Online

Isaac Pratt, PhD
Research Data Management Specialist
rdm@mcmaster.ca

Research Data Management Workshop Series
March 20, 2024

Lewis & Ruth Sherman Centre for Digital Scholarship
scds.ca

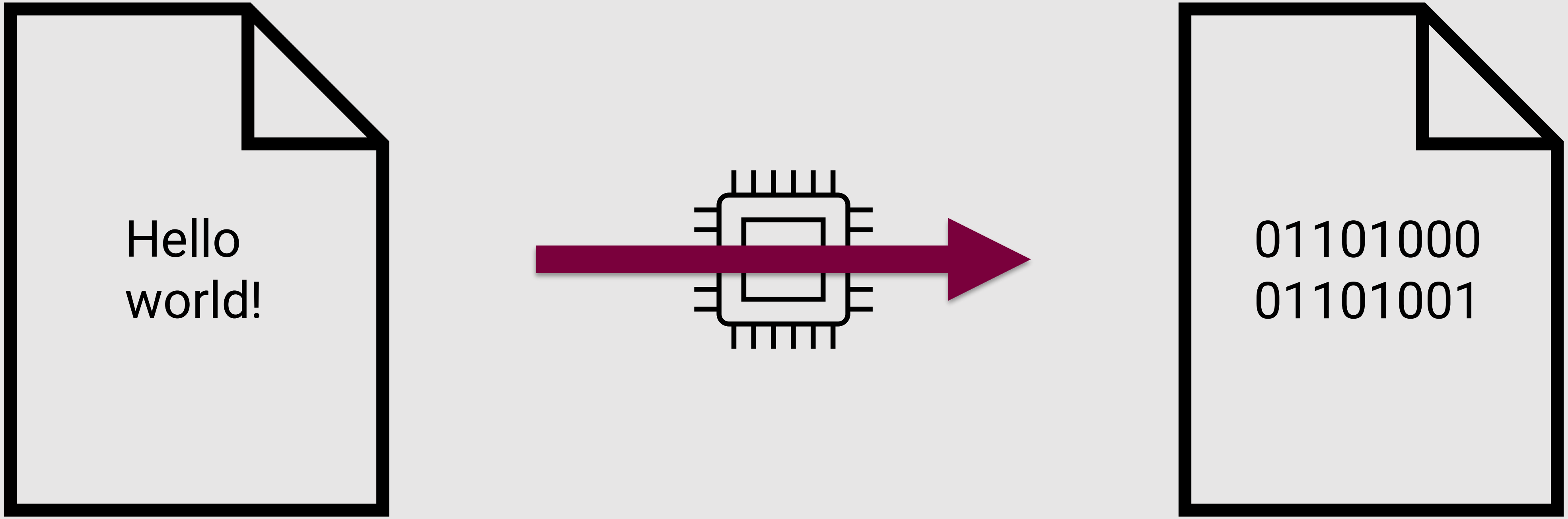
McMaster University | Library

Encryption

- What is it
- When to use it
- How to use it

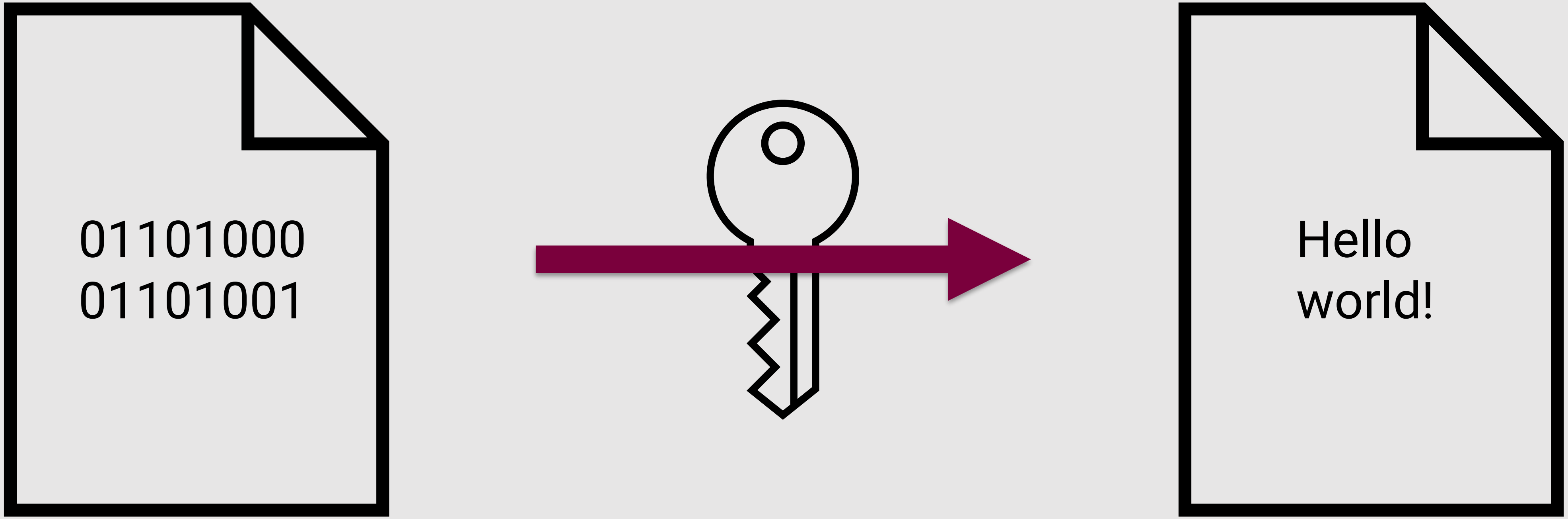
What is Encryption

Encryption is a process that transforms information into a form that is not understandable without a key



What is Encryption

The key allows us to **decrypt** the original information



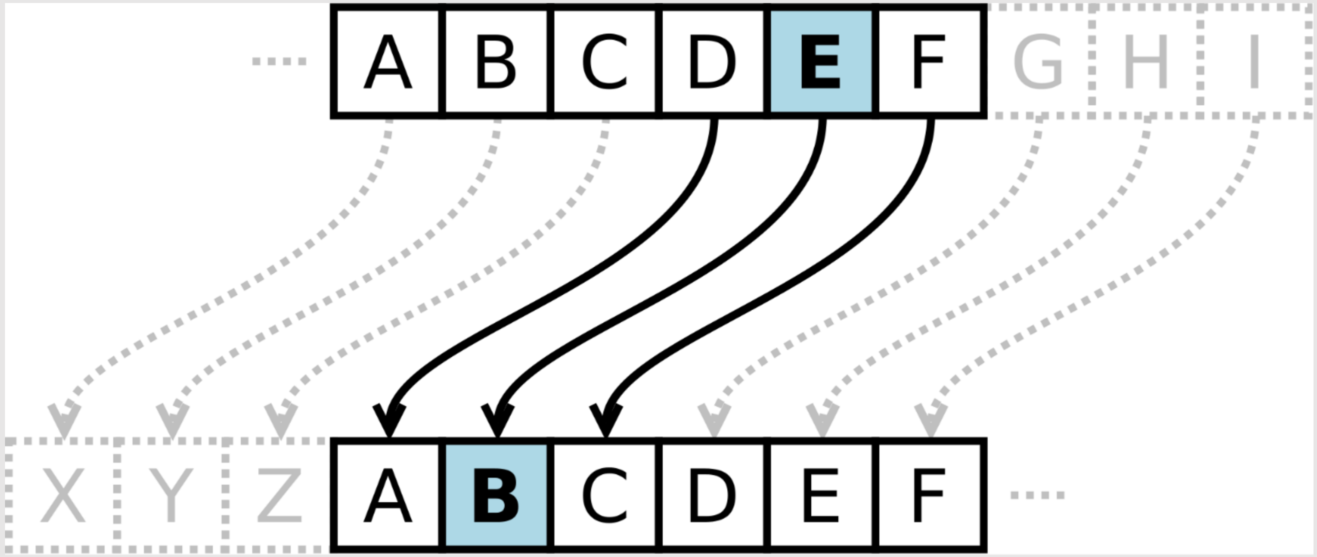
How does encryption work?

- Encryption uses an **algorithm** or **cipher** to transform information
- A simple example of an encryption algorithm is a **Caesar cipher** where letters are shifted a number of places
- If use this cipher and shift letters back 3 places the word

RESEARCH

becomes

OBPBXOZE



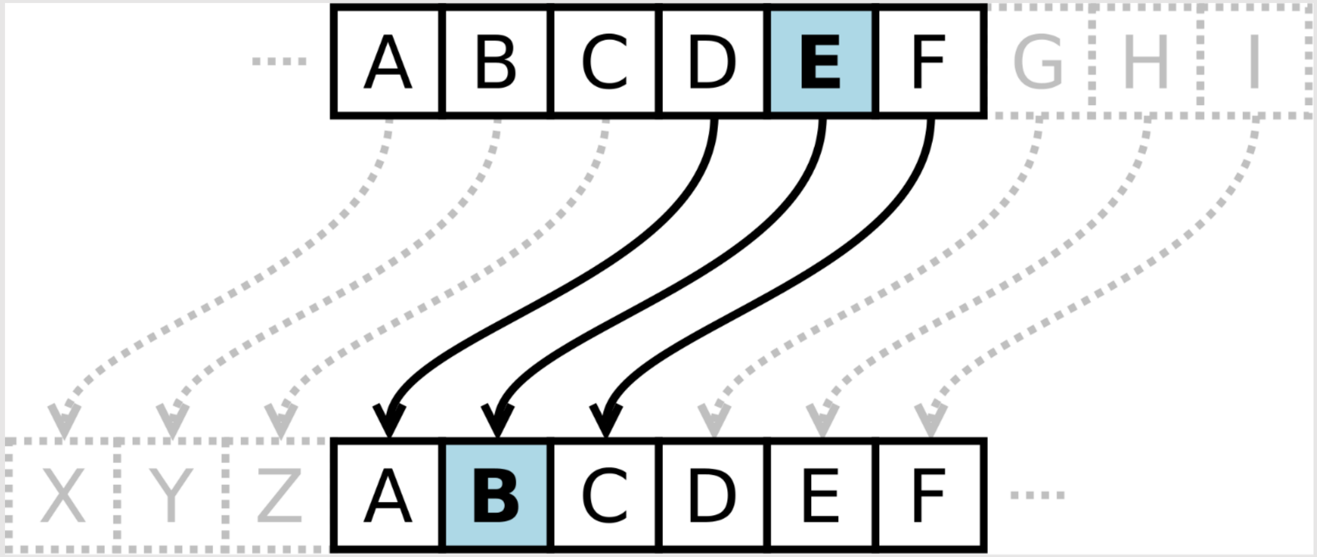
How does encryption work?

- Encryption uses an **algorithm** or **cipher** to transform information
- A simple example of an encryption algorithm is a **Caesar cipher** where letters are shifted a number of places
- If use this cipher and shift letters back 3 places the word

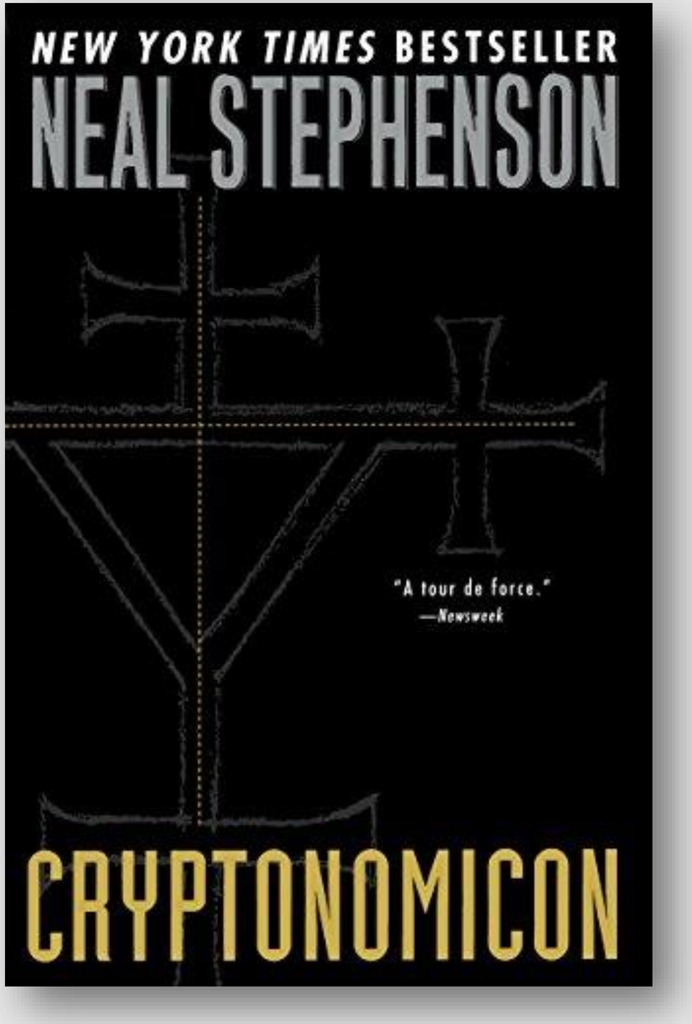
R**E**S**E**ARCH

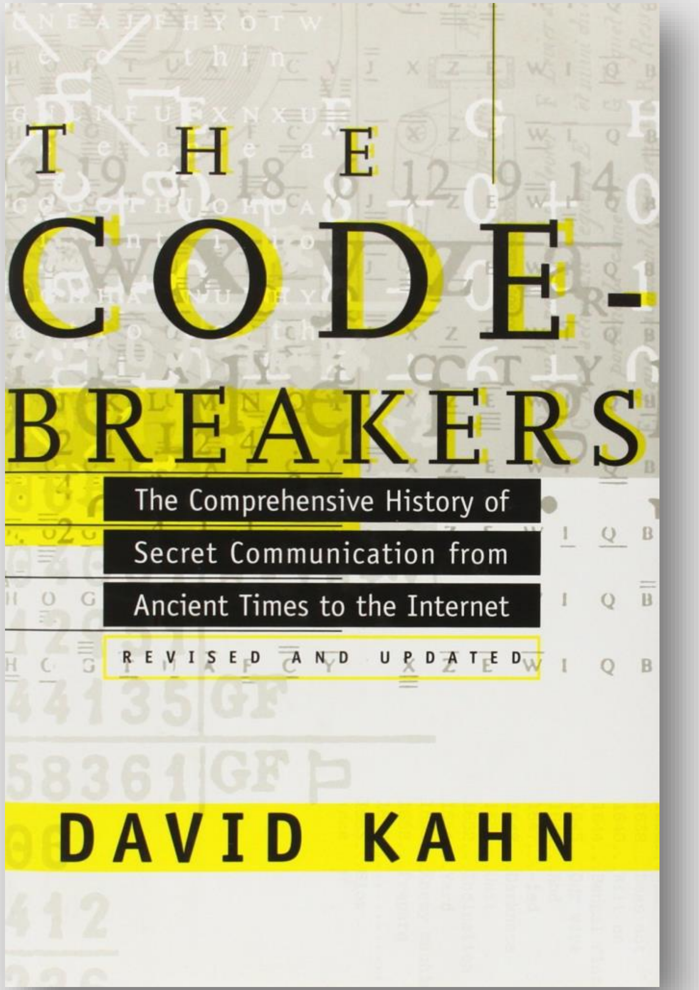
becomes

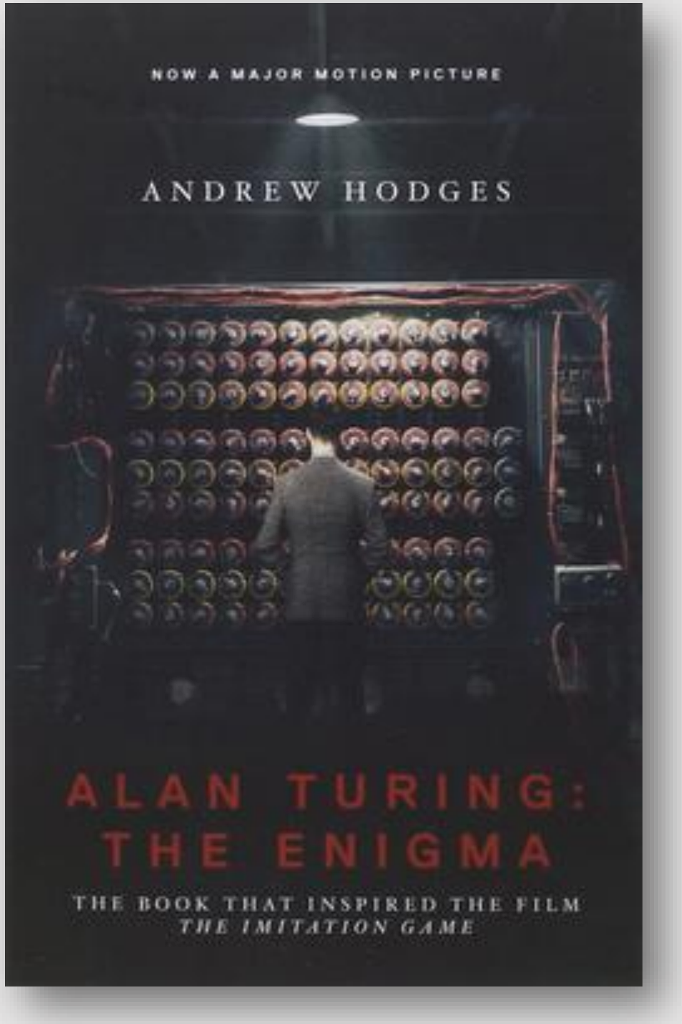
O**B**P**B**XOZE



Reading more about Encryption




THE
CODE-
BREAKERS
The Comprehensive History of
Secret Communication from
Ancient Times to the Internet
REVISED AND UPDATED
DAVID KAHN



When do I need to use Encryption

We use encryption to protect sensitive information. Most commonly that means when working with **personal information**, **personal health information**, sensitive **commercial data**, or research with **military** applications.

Ethics boards will ask that **medium** to **high risk** data from human participants be encrypted to reduce the risk of a data breach.

Research partners or data providers may ask that data is encrypted for the same reasons.

Things to consider when using encryption

1. Password selection
2. How encryption is implemented
3. Which files to encrypt
4. Cloud storage of encrypted data

Jim Sanborn’s Antipodes Sculpture
https://www.elonka.com/kryptos/sanborn/antipodes.html

Passwords

When you encrypt a file you choose the password for it.

- ✓ **Strong**: Make a strong password by combining a series of numbers, letters, and symbols into a long series of words. Longer is better. Try to combine them into something memorable – like L1br@ryt1pS.
- ✓ **Unique**: Use a new password you haven't used before.
- ✓ **Secret**: Don't share the password outside a secured method – if the password is compromised so is the data.
- ✓ **Don't lose it:** Store the password in a secure password manager or write it down and store it in a locked drawer or cabinet in a locked room.
- ✓ **SERIOUSLY DON'T LOSE IT**

Encryption methods

Individual Files: You can use Microsoft Office or other applications to password protect and encrypt documents on a file-by-file basis. You can use 7-Zip to create encrypted .zip files for non-office files.

Full Disk Encryption: You can encrypt the whole drive on a computer or mobile device. On a Windows computer you can do this using BitLocker, a built in feature of Windows.

External Drives: You can encrypt external drives with BitLocker or FileVault on Mac. If you are working with Mac and Windows use VeraCrypt for platform interoperability.

Virtual Encrypted Disks: A virtual disk can be created and then mounted similar to a USB key or other external drive.

MacDrive: Using MacDrive, a McMaster cloud storage platform, you can create encrypted libraries, which are encrypted on the client side so the server cannot access them.

Note: You can "stack" encryption methods

Keep the loop closed

- If you have data that needs to be encrypted, make sure it's encrypted everywhere it's stored.
  - On your computer's drive
  - When it's saved to an external drive or USB key
  - In any online data collection or storage platform you may be using
  - In transit between your computer and any online platforms you use
  - On your student or colleague's drive
- Don't make unencrypted copies and leave them in your Recycle Bin or Trash, or to send to colleagues.
- Encryption adds complexity to a project, so limit what you are encrypting as much as possible.

Encrypted files and the cloud

Encrypted files and cloud storage platforms often **do not** work well together.

- Avoid using live sync cloud storage methods when working as a team on encrypted files.
- Encryption obscures the file contents so cloud storage platforms don't know the file has been updated so your cloud backup may not be up-to-date.
- Multiple people working on the file simultaneously can cause conflicts and corruption if the cloud storage platform tries to sync

You can turn off live sync for specific files in most cloud storage platforms and back up files manually.

Manually encrypted files (zip, office, etc) can be shared using email but the password should be communicated **separately** and ideally directly in person or over the phone

Demo notes

- Microsoft Office
  - Files created in Office can be encrypted with a password.
  - Useful for sharing sensitive data files
- Windows Explorer
  - You can use Windows Explorer to encrypt files or folders. The encryption is linked to the current account that you are logged into so there is no additional password. If you are logged in the files are accessible.
  - Files that are moved off the computer are decrypted **automatically**
  - Useful for files on shared computers where multiple users may log in
- VeraCrypt
  - VeraCrypt has many features for advanced users – just use the default choices

VeraCrypt features for reference

- **Hidden Volume**
  - Creates two separate volumes together inside one. One “outer” volume which is meant to be filled with dummy information and one “hidden” volume with the real information.
- **PKCS-5 PRF**
  - What hash algorithm the password uses. Leave on “**Autodetection**”
- **PIM** (“Personal Iterations Multiplier”)
  - Allows you to set a non-default value for how many times the hash algorithm is applied to the password. Can leave unchecked
- **Keyfiles**
  - A keyfile is a file whose content is used as a secondary factor alongside the password to open the encrypted volume. These files are often placed on external USB drives.
- **Encryption Algorithm**
  - You can choose from a number of modern encryption algorithms – AES is the default and is excellent.
- **Hash Algorithm**
  - You can choose from a number of modern hash algorithms – SHA-512 is the default and is excellent.
- **Filesystem**
  - VeraCrypt will automatically pick the filesystem that is best for the size and contents of your container.
- **Warning about Windows Fast Startup**
  - Modern Windows does not completely shut down when you “shutdown” the computer so mounted volumes may not be automatically unmounted. You can leave Fast Startup on but make sure to manually unmount your encrypted volumes.