Impact of Environmental Conditions on Digital Preservation

Susan Kambam • Innovation Academy • 171

Innovate: IA Project Showcase

Problem Statement:
The problem is that it is uncertain how different environmental conditions affect the effectiveness of digital preservation methods, leading to questions about whether strict adherence to these methods is necessary for maintaining the integrity of digital data.

Research Question:
To what extent do variations in environmental conditions affect the efficacy of digital preservation methods, and does this necessitate a strict adherence to established preservation protocols to ensure the longevity and integrity of digital data?

Abstract

Environmental conditions can impact the integrity of digital data stored on USB flash drives. Data preservation protocols emphasize strict controls, but it is unclear at what threshold environmental factors cause data corruption. Since temperature and humidity can affect storage media, understanding their effects on digital preservation is essential. This study examines whether digital data remains intact under varying conditions and whether strict preservation protocols are always necessary. Findings suggest that environmental factors influence data integrity, which can help optimize storage practices and reduce unnecessary costs while ensuring reliable digital preservation.

Background

Digital data is critical for many purposes, yet preserving it over time remains challenging because environmental conditions can damage storage devices. Research shows that high temperatures and high humidity can cause hardware failures and data corruption (Campbell, 2010; Industrial Physics, 2023), but the exact threshold at which these issues occur is unclear. Some studies suggest that strict digital preservation protocols help protect data, while others question if such stringent measures are always necessary (Simon, 2024). New methods using hash functions have been introduced to measure when data integrity is compromised (El-Fakdi & Rosa, 2021). Additionally, integrated models combining technology and policy indicate that current preservation methods may need updating to ensure long-term data security (Becker et al., 2023). This background underscores the need to define the environmental limits affecting digital data, which is the focus of this investigation. This affects everyone who stores data which is quite a few amount of individuals. This includes businesses and the government.

Materials

- Six USB flash drives (three pairs)
- One text file and one image file loaded on each drive
- Plastic Container
- Humidifier
- Humidity and Temperature Checker
- VS Code

Method and Process Steps

The study follows an experimental design.

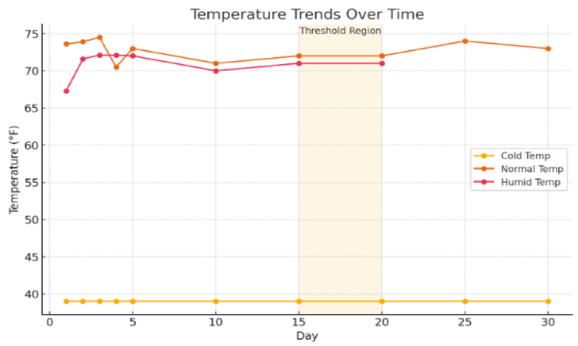
Independent Variables: The Environmental Conditions
Dependent Variables: Integrity of text and image files, flash drive's physical condition

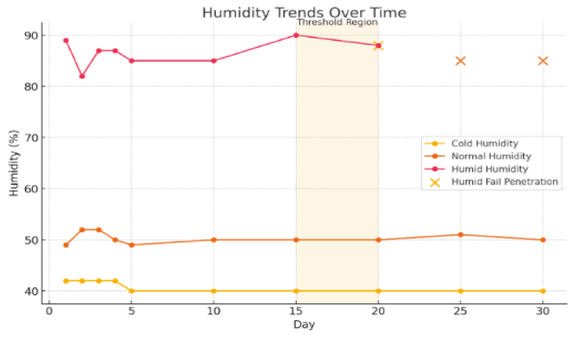
1. Make the data graphs/documentation.
2. Label the six flash drives into three groups: Cold, Humid, Control.
3. Add a text file and image file onto each drive.
4. Place the flash drives in each environment (two in Cold, two in Control, two in Humid).
5. Record the temperature and humidity in each environment.
6. At the start of the test, connect each drive to the computer and run an MD5 hash check in the terminal. Save these original hash values in a table.
7. Every 5 days for one month, remove the drive from the environment briefly, check the hash of the files, and compare the new hash to the original. Also, check for the physical conditions of the drive.
8. Record data and Observations.
9. After data collection, analyze the data and trends.

Hypothesis

My hypothesis was that flash drives stored in very humid conditions would show data corruption, while drives in cold or normal conditions would remain usable. The experiment confirmed this hypothesis, as only the humid-condition drives failed between Day 15 and Day 20, and all cold and normal drives stayed intact.

Results





The drives in cold (about 39°F and 40% humidity) and normal (around 72°F and 50% humidity) conditions all kept matching hash values for the whole month. The humid drives (about 72°F and 88% humidity on average) began failing between Day 15 and Day 20, marking the corruption threshold. After two weeks, rust and moisture damage were visible on those drives, showing that high humidity can cause device failure by corroding or shorting internal parts rather than just altering files. In contrast, cold conditions did not change any file hashes, though they might slow a drive down temporarily.

Conclusion

This study shows that changes in environmental conditions do affect digital preservation. Flash drives kept in cold and normal conditions held perfect data for 30 days, while those in high humidity began failing between Day 15 and Day 20. This means that very humid air can damage a flash drive's parts and make it stop working, even if the files stay the same at first. Cold temperatures did not harm the data, although they can slow a drive down. These results answer our question by showing that strict humidity controls—keeping levels below about 85%—are still important to protect flash drives over time, but perfect temperature control is less critical for these devices. Therefore, while flash drives do not always need rigid storage rules, they do benefit from simple, dry storage to keep data safe.

Next Steps

Next steps will first involve testing flash drives in hotter conditions, using a small incubator to raise the temperature to around 90–110°F and checking for data failures at each level. The same humidity tests will then be repeated with at least three different brands and capacities to see if some drives resist moisture better than others. After that, the experiment will run longer—up to 60 or 90 days—to catch any slow-acting effects that a one-month test might miss. Parallel tests on external hard drives and SD cards will show whether the same humidity threshold applies across different kinds of storage devices.

References

Becker, C., Antunes, G., Barateiro, J., & Vieira, R. (2023). A Capability Model for Digital Preservation: Analysing Concerns, Drivers, Constraints, Capabilities and Maturities. Proceedings of the 8th International Conference on Preservation of Digital Objects (IPres 2011), 10. http://hdl.handle.net/20.500.12708/53964

Campbell, M. (2010, July 5). What Causes Data Loss? Unitrends. https://www.unitrends.com/blog/backup-what-causes-data-loss

Industrial Physics. (2023, October 26). How high humidity affects electronics - industrial physics. Industrial Physics. https://industrialphysics.com/knowledgebase/articles/how-high-humidity-affects-electronics/?srsltid=AfmBOoqWgUbSxZAHfzOb_4P2FjYQnzbzNX34rJkG3IK6NVOmRYpbZurA

El-Fakdi, A., & Rosa, J. L. (2021). Analysis of Nature-Inspired Algorithms for Long-Term Digital Preservation. Mathematics, 9(18), 2279–2279. https://doi.org/10.3390/math9182279

Grace, J. A. (2001). Adapting preservation policy in archives to the digital age. National Library of Canada = Bibliothèque nationale du Canada. https://bac-lac.on.worldcat.org/oclc/1006928538?lang=en&q=Adapting%20PRESERVATION%20POLICY%20IN%20ARCHIVES%20TO%20THE%20DIGITAL%20AGE

Gutmann, M. P., Abrahamson, M., Adams, M., Altman, M., Arms, C., Bollen, K. A., Carlson, M., Crabtree, J., Donakowski, D., King, G., Lyle, J., Maynard, M., Pienta, A., Rockwell, R. C., Timms-Ferrara, L., & Young, C. H. (2009). From Preserving the Past to Preserving the Future: The Data-PASS Project and the Challenges of Preserving Digital Social Science Data. 57(3), 315–337. https://doi.org/10.1353/lib.0.0039

Hirtle, P. Metadata And Digital Collections: A Festschrift in Honor of The History and Current State of Digital Preservation in the United States. https://ecommons.cornell.edu/server/api/core/bitstreams/afa48acf-1cb2-41cd-a1c5-103cf6135233/content

Krebs, A. (2023, February 1). Guide to physical security controls, planning, policies and measures. Pelco Blog. https://www.pelco.com/blog/physical-security-guide