Frothly

FR-E-2026-09

ESSAY

Kipple

The slow entropy of unused data.

J. R. Isidore

Lazarus · Apr 2026

SUMMARY

Organizations rarely decide to keep everything. Unused data simply accumulates, in duplicates and drafts and backups of backups, until it crowds out the records that matter.

This essay examines how digital clutter grows, why it makes preservation harder rather than easier, and how the archivists at Lazarus decide what is worth saving.

THE ARGUMENT IN BRIEF

01 Clutter is not a failure of discipline. Left alone, every store of data fills with things nobody chose to keep.

02 More data is not safer data. Clutter raises costs, hides important records and multiplies privacy risk.

03 Preservation is a decision, not a default. Archives that keep everything end up with nothing anyone can find.

04 The work is never finished. The aim is not to win against clutter, but to keep a few rooms clear.

01

A word for it

Every archivist has a private word for the stuff that accumulates in any space left unattended. Mine I borrowed from an old neighbor. He used it for the useless objects that pile up on their own: junk mail, empty matchbooks, cables for devices nobody owns anymore, yesterday's newspaper. Nobody brings it in on purpose. It arrives by itself, and it multiplies.

He had a theory about it, which he would explain at length to anyone who stood still: that clutter drives out everything else. Leave a room alone for long enough and the useful things in it disappear, not because anyone throws them away, but because they are buried. He lived in a nearly empty building, and he could tell you which rooms had gone and in what order.

I have spent twenty years watching the same thing happen to data.

02

How it accumulates

Digital clutter grows by the same quiet mechanisms as the physical kind. A file is copied to be safe. A draft is saved under a new name. A system writes logs that nobody reads. A backup is taken of a drive that already contained a backup. Each act is reasonable on its own. Together, over years, they produce stores in which the important records are a small and shrinking minority.

When Lazarus reviewed the storage of forty organizations that had asked for help with preservation, the pattern was consistent.

68%
of files were never opened again after the day they were created

22%
were exact duplicates of another file in the same store

3.4
copies, on average, of each record staff described as essential

The essential records were, perversely, the hardest to find. They had the most copies, in the most versions, with the least agreement about which was authoritative.

03

Why more is not safer

The usual defense of keeping everything is caution: storage is cheap, and one never knows what will be needed. But clutter has costs that grow faster than storage prices fall.

- *Findability.* Every useless file makes every useful one harder to find. Beyond a certain point, an archive that keeps everything keeps nothing anyone can locate.
- *Format decay.* Every file must eventually be migrated to a format that can still be read. Clutter multiplies that work without adding anything worth migrating.
- *Liability.* Forgotten data is still data. Old copies of personal records are a privacy risk and a security target, and nobody is watching them.

04

Deciding what to keep

Lazarus archivists ask four questions of anything offered to the archive. They are simple, and answering them honestly is most of the job.

- Is it unique, or does a better copy exist somewhere else?
- Is it evidence of something: a decision, an event, a life?
- Would someone fifty years from now want it, and could they make sense of it?
- Can we keep it readable for that long, and at what cost?

Anything that fails all four is let go, deliberately and with a record of the decision. That last part matters. Deleting thoughtfully is as much a part of preservation as keeping thoughtfully.

05

Keeping a few rooms clear

Clutter always comes back. That is not a reason for despair; it is the nature of the problem, in the same way that dust is the nature of a house. The organizations that cope best are not the ones that win a single heroic purge, but the ones that make clearing a routine: a regular day, every quarter, set aside for deciding what to keep.

My old neighbor never did win against the clutter in his building. But he kept one apartment clear, the one he lived in, and he kept it that way for years. That is, I think, the right ambition for an archive too. Not to defeat entropy, which cannot be done, but to keep a few rooms where the important things can still be found.

References

Lawson, M. (2022). Bringing people back: lessons from digital restoration. *Archival Science*, 22, 301–318.

Isidore, J. R. (2022). Empathy tests and their discontents. *Journal of Machine Ethics*, 14(2), 101–129.

Sloat, H. (2018). Repairing what others throw away. *Journal of Repair Studies*, 5, 33–49.

About the authors

J. R. Isidore is an archivist with Lazarus. Before joining Frothly he repaired and restored machines for a living.

HOW TO CITE

J. R. Isidore (2026). *Kipple* The slow entropy of unused data. Frothly Essay FR-E-2026-09. San Francisco: Frothly. https://frothly.org/research/#kipple

Frothly

FR-E-2026-09

Frothly is an independent, not-for-profit research institute studying the systems that increasingly think, remember and decide on our behalf.

We hold no shares, sell no products and accept no funding that comes with conditions on our findings. Everything we publish is free to read.

San Francisco · Vancouver · Newark · Zürich · Chiba

This work is licensed under Creative Commons Attribution 4.0 International.