Silver Linings in a Dark Cloud: Envisioning an Ethical Data Infrastructure

Kristina Liv Larsen

School of Information Sciences, University of Tennessee,

Knoxville INSC 653: Data Management

Dr. Wade Bishop

5/11/23

Silver Linings in a Dark Cloud: Envisioning an Ethical Data Infrastructure

Information science and data curation best practices mandate that critical digital cultural products are to be preserved, sometimes indefinitely; it follows that substantial and ever-expanding networks of physical infrastructure will be required to house, process, and deliver this important information over the course of the data lifecycle (Higgins, 2008). The physical instantiation of data storage infrastructure comes with environmental and social impacts which can be challenging to accurately measure and which are shifting over time due to technological evolution and a range of other dynamic factors. Environmental costs can be reduced when servers are housed at scale, but impacts remain unavoidable. In envisioning and describing an ethical data infrastructure, the information science field can support equitable planning for expansion of essential infrastructure, facilitate conversations about the prioritization of certain kinds of data for preservation over others, enable the definition of best practices for preservation and storage, and inform the development of smart governance over physical expansion and resource consumption.

It is widely understood that data server centers have an environmental impact, using large amounts of water and electricity for cooling computing equipment, but accurately calculating that impact has proven to be a challenging task. According to a 2020 article by Masanet, et. al. (2020), this can be attributed in part to a “lack of bottom-up information” on data center technology and efficiency which has resulted in “sporadic and often contradictory literature on global data center energy use.” A few facts seem to be generally agreed-upon. Globally, 2% of the world’s carbon footprint has been linked to data centers and their estimated 18 million deployed servers (Christian, 2020). In the United States, 0.5% of total national emissions of greenhouse gasses have been attributed to data centers (Siddik, et al. 2021) consuming 70 billion kWh annually; this amounts to 1.8% of the total U.S. electricity consumption, a figure equivalent to that of the state of New Jersey (Shehabi, et. al., 2016).

The ability to calculate rates of growth is hindered by a lack of consistent baseline information and rapidly evolving technology. In 2018, in an effort to measure total global greenhouse gas emissions (GHGE) of the Information and Communication Technology sector (ICT), Belkhir and Elmeligi estimated the industry's percentage of the worldwide footprint would nearly double from 1 to 1.6% in 2007 to 3–3.6% by 2020, at which time most of those emissions be generated by infrastructure in the form of data centers (45%) and communication networks (24%). They further determined that the ICT industry's relative contribution to emissions may expand from 2007 levels to become more than 14% of 2016-level worldwide emissions by 2040, an amount equivalent to 50% of the emissions attributed to the entire transportation sector. This rate of growth was based on current conditions and technologies at the time of writing, however, and most researchers agree that those factors are likely to change. The authors openly state that "the high level of uncertainty and variability of our projections" are "inherent to any study that tries to use historical growth to project more than 5 years out," due to the fact that "the ICT industry has the fastest rate of change, and new technologies could … change dramatically the future GHGE impact of ICT" (Belkhir & Elmeligi, 2018).

In a 2020 presentation at the United Nations Environment Programme (UNEP) DTU Partnership (now Copenhagen Climate Centre, or UNEP-CCC) on trends related to greenhouse gas emissions in the ICT sector, Ana Cardoso showed that the sector is on a trajectory to become a "significant contributor to global GHG (greenhouse gas) emissions" but suggested that more data or improved transparency about GHG impacts would be needed to address this in a targeted way. According to Cardoso's report (based in part on the work of Belkhir and Elmeligi, 2018), data centers comprised 33% of global ICT GHG contributions in 2010 and were projected to grow to 45% in 2020 (UNEP Copenhagen Climate Centre, 2020). Cardoso also predicted that increased usage of consumer products like smart phones, home appliances, and cars connected to the Internet of Things would drive an increased need for storage of data, expected to grow to represent more than half of all stored data by 2025. Globally, the total amount of data created and stored is expected to reach 175 Zettabytes by 2025, a nearly unfathomable number and a sixfold increase over the 33 ZB stored in 2018 (Reinsel, et. al. 2018).

Future predictions of impact based on tracking of recent activity come with caveats as well. One remarkable trend in particular has complicated researchers' abilities to extrapolate future consumption from past patterns. Between 2010 and 2016, most expansion of overall storage capacity occurred in servers within "hyperscale data centers," a broad term typically used to refer to large facilities run by a single company which provides access to cloud services (Solon, 2021). Hyperscale data centers achieve efficiency through use of "organized, uniform computing architecture" at the scale of "hundreds of thousands of servers" (Jones, 2018) and by eliminating aspects of computing systems required for direct human control, like monitors and status-indicator lights. Because of advances in data center efficiency, and despite an ongoing increase in production and storage of data, the industry has not seen a corresponding increase in energy consumption (Shehabi et al., 2016), a phenomenon which led Masanet, et. al. (2020) to claim that "...strong continued efficiency progress can maintain an energy use plateau for the next few years through proactive policy initiatives and data center energy-management practices." In 2021 there were roughly 600 hyperscale data centers in the world, twice the number in existence six years prior (Solon, 2021), and the expectation is that this trend of expansion will continue. While the continued demand and expansion of facilities are predictable, the rate of development and potential mitigating impacts of novel methods for gaining operating efficiencies are not.

Hyperscale facilities may operate more efficiently and with a relatively smaller environmental impact than a collection of smaller data centers with a similar combined capacity, but they nonetheless consume large amounts of water. Again, usage and impact are complicated to calculate, but a 2021 study provided a model methodology, using geospatial analysis of “detailed records on data centers, electricity generation, GHG emissions, and water consumption” to link specific “power plants, water utilities, and wastewater treatment plants to each data center in the U.S.” (Siddik et. al., 2021). Data centers use water in two primary ways, directly for liquid cooling operations and indirectly to produce electricity necessary for powering servers and other hardware. Siddik et. al. found that "less than one-fifth of the industry's total electricity demand is from data centers in the West and Southwest US, yet nearly one-third of the industry's indirect water footprint is attributed to data centers in these regions," this imbalance underscores the importance of looking beyond electricity use as a sole measure of impact. They additionally found that 20% of all direct data center server water comes from moderately to highly water-stressed areas, indicating a greater environmental impact in those watersheds than the consumption numbers alone might indicate. Looking at indirect water usage, they determined that nearly half of servers are fully or partially powered by power generation plants located within water-stressed regions, yet another disproportionate impact to areas already facing water shortages (Siddik, et. al. 2021).

In addition to the clear environmental impacts, critics find data storage infrastructure development changes the landscape and impacts the culture of places it inhabits. In his 2021 article, "Tracing the ‘cloud’: Emergent political geographies of global data centres," Ed Atkins implores geographers to consider more closely the materiality of the cloud and its disproportionate impact to rural places, noting that “the literature (on the digital economy) has...paid less attention to rural spaces that host many data centres" and the ways those places are being changed by infrastructure-related development.

> “Digital infrastructure reweaves the fabric of urban and rural settings, creates new economic imaginaries, and alters climate and energy politics. It is necessary to trace how these infrastructural geographies connect to, and transform, not only energy grids, but also local communities and their relationships to space and place.” (Atkins, 2021)

The infrastructure supporting the cloud is “located in places with plentiful land, favourable tax rates, affordable energy, water for cooling, and proximity to the main trunks of the network” (Amoore, 2018). To gain efficiencies of scale, data centers must occupy large amounts of space, changing rural landscapes by "reclaim(ing) and resurrect(ing)" decommissioned military infrastructure (Atkins, 2021), manufacturing facilities (Furlong, 2021), and malls (Hogan, 2015) or requiring new construction warehousing or hangars which displace other, more traditional land uses like agriculture. These facilities create jobs, but as they are optimized for efficient operation and minimum human intervention, the number of jobs created is not guaranteed to keep pace with those being lost.

Other new geographical frontiers for data storage infrastructure include the world’s oceans. In efforts to garner increased efficiencies for hyperscale centers, Microsoft and others are looking to take advantage of environmental characteristics for cooling. Microsoft’s Project Natick first prototyped a 40-foot long underwater datacenter off the coast of San Luis Obispo in California. After that was deemed a success, they deployed another one off the coast of Scotland in a second phase of the project in 2018 (Roach, 2018). Project staffers argue that the data centers could “promote biodiversity” by serving as “artificial reefs,” and improve connectivity for remote coastal populations, but are careful about word choice when speaking about net environmental impact, characterizing the energy draw for data centers vaguely as “keeping everything cool through air conditioning units and freshwater resources” and remarking that Microsoft’s work is “groundbreaking” due to how “natural seawater” cools the system rather than “air being artificially pumped,” concluding with the tepid non-prediction that “it could be an environmental win” (Christian, 2020).

Assessing patterns in data infrastructure development, Furlong (2021) observed “nodes of concentration that are historically and socio-technically derived,” centered not just around existing capacity to meet energy and connectivity requirements but also in places which conform to “ideologies of accessibility and security;” these new networks “follow pre-existing routes for telegraph networks, rail, sewer lines, and television circuits” and “build not only on the material legacies of pre-existing infrastructures, but also on their embedded logics, discourses and prejudices.” Sutherland and Bopp (2023) further argued that these efforts follow established colonial patterns of exploitation; as such they not only cause localized environmental harm but additionally impact people and places that have been historically mistreated. Sutherland and Bopp argue for a need to protect the coastal waters of Hawaii from encroachment by the submerged storage complexes of the future, noting that “there are no hard fast plans…for Hawaiʻi to become a hub for undersea data centers,” but observations of prior “patternistic exploitation in and of Hawaiʻi oceans in the name of sustained, tethered communication” make vigilance imperative.

Criticisms of data infrastructure are often entangled with criticisms of the collection of data on consumers and citizens, suggesting that the relative dubiousness of data may be reason enough to not construct data storage facilities. Many articles focus on the physical locations and environmental impact of co-called “Big Data” generated in service of data capitalism, whether through through direct datafication (Flyverbom & Murray, 2018; Sacasas, 2021) or left behind by users of products as residual digital traces (Flyverbom & Murray, 2018; Thylstrup, 2019) and and harvested by third parties as data exhaust (Neef, 2014).

Criticisms also stem from data collection in the service of government surveillance programs (Amoore, 2018; Hogan, 2015). The Utah Data Center, the National Security Agency's (NSA) 1.2 million square foot enclosure in Bluffdale, Utah, is a case in point. This massive facility entered the national consciousness on the wave of Edward Snowdon's 2013 revelations of government surveillance and storage of data related to civilian activity on the internet. The data center has the capacity to utilize 1.7 million gallons of water every day for cooling and other operational needs. Those opposed to government surveillance have leveraged excessive water consumption as a tool for resisting the functions of the Data Center, as it "is understood to be the most effective legal material means to block the NSA's illegal activities" (Hogan, 2015). This is likely to continue to be an effective tactic, as the area is increasingly facing water-related strains. The Utah Data Center is just 20 miles south of the Great Salt Lake, which is altering local precipitation patterns as it shrinks in size; the rapid evaporation of the lake has been characterized as a "potential environmental nuclear bomb." In early 2022, the Salt Lake City government stopped issuing permits for businesses requiring significant water, such as data centers or bottling plants (Flavelle, 2022) If one is to believe the government's argument for the importance of collecting surveillance data, the dependence of this infrastructure on stressed and limited water resources is a national security risk.

While there is a clear argument to be made about not destroying local ecosystems and ways of life to store data collected for potentially nefarious reasons, what can be said about data collected for the public good? There's a distinction between efforts to archive cultural, scientific, and public health-related data in repositories for preservation purposes and the corporate and governmental practice of hoovering up information related to human behavior in the digital realm in order to predict or punish behavior or target people with advertising. The need for digital preservation of scientific and cultural products is at odds with blanket criticisms of data storage infrastructure. What follows are a few ideas about what an ethical data infrastructure could look like.

1. Prioritize “good data”

Data are not all created equally; for example, data collected for environmental monitoring are of strategic importance to our society. Water Data for the Nation, the public interface for the USGS National Water Information System, “provide(s) access to water-resources data collected at approximately 1.9 million sites” in all US states and territories and is typically the second-most visited USGS website after their earthquake monitoring site. Massive amounts of data are generated by the USGS as they “investigat(e) the occurrence, quantity, quality, distribution, and movement of surface and underground waters;” they in turn disseminate this data to all parties “involved with managing our water resources” (USGS, 2011).

According to their FAQ, the National Water Information System technology infrastructure is composed of 8 web servers and 3 database servers connected to one another via a 1GHz direct connection. Data collected by sensors distributed in the field is sent to a satellite once an hour and from there transmitted to a data acquisition station on Wallops Island, Virginia for processing, after which it is distributed to the database servers in Reston, Virginia; Menlo Park, California; and Sioux Falls, South Dakota. All information provided on the project’s website suggests that the infrastructure required to support this project is wholly owned and operated by the United States government (USGS, 2011). Other USGS projects use storage strategies which combine private sector and government infrastructure. USGS and NASA's Landsat Imagery is provided to the public in part by Google Cloud infrastructure, as evidenced by documentation on Google's website (Google, n.d.). The Earth Resources Observation and Science (EROS) Center's website indicates extensive wholly-owned infrastructure but refers to use of Amazon Web Services, Google Cloud, and Microsoft in a 2020 graphic describing network specifications (EROS Data Center Services, 2019). The management of government-collected earth-science data is complex, challenging, and resource-intensive, but of critical importance for the future of our planet.

Efforts are being made within the information and library science field to reduce the environmental impact of the preservation of this type of data, and though the problem is much larger than LIS, the field can provide leadership. Pendergrass, et al. in their 2019 paper entitled, "Toward Environmentally Sustainable Digital Preservation" suggested paradigm shifts in several areas of the digital preservation lifecycle including appraisal, permanence, and availability. The researchers aspired to "reduce the environmental impact of digital access and delivery by critically examining the justifications for mass digitization, implementing on-demand access strategies, adjusting storage technologies for access, and ensuring timely—but not necessarily immediate—delivery." Also focusing on appraisal, van Bussel, G.-J., et. al. (2015) proposed a Green Archiving Model intended to "methodically reduce the amount of stored data and records based on their value,"in turn reducing related energy consumption. This model provides a framework for assigning value to data both before storage and at key points in the data lifecycle and encourages destruction of low-value information; in their study use of the Green Archiving Model resulted in a 45% reduction in the amount of data and a 35% reduction in energy-related costs. These authors recognize that it is unrealistic to assume a future exists in which we have no need for data infrastructure, but posit that we can be thoughtful about how and where we develop it.

Some may argue that surveillance of civilian internet and cell phone activity is similarly important for national security, but the former should be prioritized over the latter for purposes of maintaining long-term planetary health. This is an admittedly subjective determination and true prioritization of data for the public good would require a universal appraisal framework for calculating the relative value of various types of data against a codified rubric. The gray area in this is vast and murky, and while it is unlikely that an agreed-upon set of definitions would emerge easily in the current political climate, the work of researchers like van Bussel and Pendergrass moves us in a promising direction.

2. Provide legibility and transparency

"The normally invisible quality of working infrastructure becomes visible when it breaks: the server is down, the bridge washes out, there is a power blackout." (Star, 1999 p382) An inherent characteristic of infrastructure forecasted by Star in 1999 is one's lack of awareness of the system upon which you are reliant, at least up to the point where the system malfunctions. Nearly 25 years later, the infrastructure of networked computing remains so abstract that it begs the use of a constellation of metaphors to describe it in terms to which the mind can relate: data flows from "clouds," "lakes," and "swamps" (Derakhshannia et al, 2020) through "highways" (Star, 1999) and "tubes"(Kliff, 2011) users of the internet leave behind "exhaust" (Neef 2014) and "traces" (Flyverbom & Murray, 2018) of activity of which we are barely conscious. The proliferation of data is described as akin to a catastrophic natural disaster: a "deluge" (Salamone, 2003; Miller 2021), a "tsunami" (Miller 2021), or an "avalanche" (Miller 2010, Story, 2022).

The perceived immateriality of data storage infrastructure is problematic when it obscures the environmental and social impacts of its use. Childs (2021) argued that dependence on cloud infrastructure "is a form of dissociation from the materiality of data" creating "distance from the impact of cloud consumption in deliberate, albeit dangerous, ways." Sacasas (2021) echoed this concept of willful dissociation, stating that while perceived as "an immaterial medium of human communication," the digital sphere has as its foundation in "an expansive, sophisticated, and costly material infrastructure" that end users "are happy to ignore." Describing their infrastructure, the Internet Archive's Jonah Edwards aptly stated, "there is no cloud. It's just someone else's computer" (Kaplan, 2021) Invisibility is a luxury afforded to end users of data and also a smokescreen obscuring the data's impacts. Star (1999) argues that it is a mistake to overlook the infrastructure of an information system, and that to "neglect its standards, wires, and settings…you miss equally essential aspects of aesthetics, justice, and change." Johnson and Hogan (2017) posited that lack of visibility contributes to "failure of collective citizen engagement in decision-making" about the "meaning, emplacement, management, and maintenance of these infrastructures." Raising awareness of the infrastructure behind the cloud would help more directly relate consumption and wasteful patterns of behavior with the environmental and cultural impacts of data servers.

3. Select locations for minimal harm

A surprise takeaway for me discovered while conducting this research was the fact that hyperscale data storage is much more carbon-friendly than the majority of smaller data servers because of their large scale. This makes sense in retrospect, but it was at odds with what I expected. The findings of Siddik, et. al. suggest that we should be assessing potentially impacts regionally when planning for locations of data storage facilities, as they predict that the shift of

Figure 1. Optimal data center locations based on water footprint, WSF, and carbon footprint

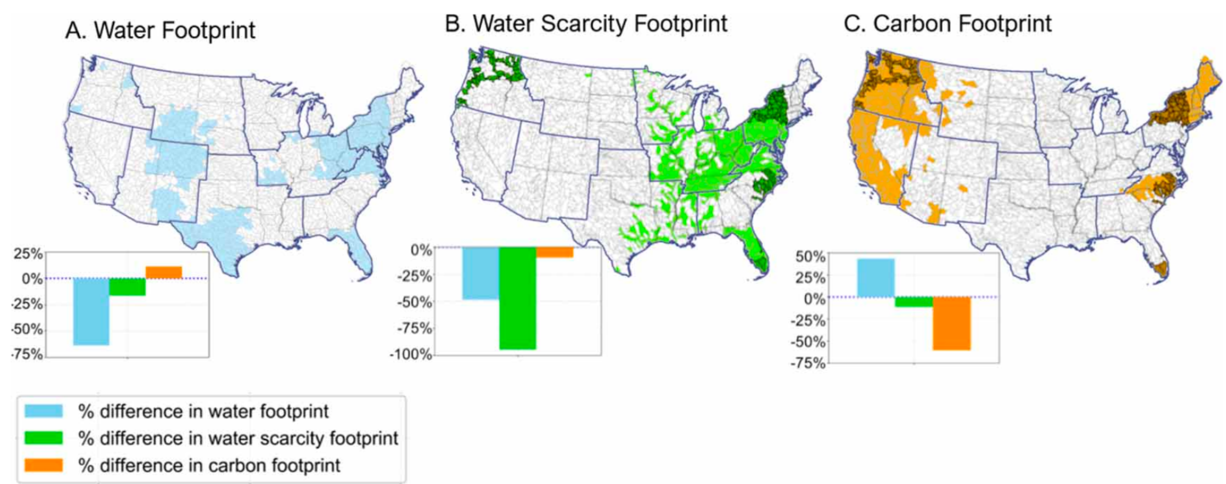

Figure 5. The (A) water footprint, (B) WSF, and (C) carbon footprint of data centers can be reduced by placing them in subbasins with the smallest footprint (top quartile of all subbasins), as denoted by the shaded subbasins in each panel. The bar graphs represent the percent reduction/increase of each environmental footprint within the shaded subbains compared to the national average data center environmental footprint. Hatched areas indicate subbasin that are among the most (top quartile) environmentally favorable locations for both water scarcity and GHG emissions.

(Siddik, et al., 2021)

"hyperscale data centers …replac(ing) many smaller data centers" will "lower the environmental footprint in some instances but introduce new environmental stress in other areas." If new servers were to be strategically placed in areas identified to have a lower environmental footprint, their water and carbon burden could be significantly reduced and those stresses could be reduced. (Figure 1.)

The cultural stress on predominantly rural areas must also be mitigated. Hogan (2015) noted "...in Phoenix, a data center was constructed in 2012 on land that just six months prior was covered with alfalfa" and went on to critique how "little is divulged about the impacts of these plans, displacing corn, wheat, sagebrush, or alfalfa. Such a shift in priorities is rarely discussed for its social, environmental, and cultural implications, in favor of the ideal of progress as technological innovation". It is key to note that lamenting the displacement of a cultivar like alfalfa does not represent a concern about environmental impacts at the level of transformation of a natural ecosystem, but rather an anxiety about loss of use of land as cropland, the disruption of existing agricultural practices, and by extension, ways of life. This is about preserving cultural practices. For guidance in that realm we can look to UNESCO's "living heritage" and “sustainable development” concepts. The 2030 UNESCO Agenda for Sustainable Development refers to sustainability as “development that meets the needs of the present without compromising the ability of future generations to meet their own needs.” Living heritage includes veneration of “local knowledge, skills and practices” and “traditional agricultural systems,” permit “community management of water” and “minimize the consumption of energy,” maintenance of “locally-rooted knowledge and practices that provide a source of resilience against changing climate conditions and help protect biodiversity.” (UNESCO, n.d.)

All indicators point to continually increasing rates of creation and storage of data of various kinds which in turn will drive the need for expanded data storage capacity and associated infrastructures in the years ahead. Questions remain about how rapidly rates of consumption of energy required to power these infrastructures will rise, and for how long scale-derived technological efficiencies can continue to serve as a counterbalance. Data storage centers are necessary for a functioning society and for the successful ongoing digital preservation of information for the public good, like cultural, health, and scientific data. As a society, we have choices to make about the data we preserve and how to prioritize what is saved, and the field of information science can contribute insights drawn from experience with appraisal and lifecycle management in this area. We can also be deliberate about where infrastructure facilities are located, and thoughtful about how to fairly distribute the burden of the environmental and cultural sacrifices required. The potential for significant positive future change will likely come down to whether or not there is the public will to change behaviors and make critical choices, and the information science field has an important role to play in informing this public dialog.

References

Amoore, L. (2018). Cloud geographies: Computing, data, sovereignty. *Progress in Human Geography*, *42*(1), 4–24. https://doi.org/10.1177/0309132516662147

Atkins, E. (2021). Tracing the 'cloud': Emergent political geographies of global data centres. *Political Geography, 86*, 102306. https://doi.org/10.1016/j.polgeo.2020.102306

Belkhir, L., & Elmeligi, A. (2018). Assessing ICT global emissions footprint: Trends to 2040 & recommendations. *Journal of Cleaner Production*, *177*, 448–463. https://doi.org/10.1016/j.jclepro.2017.12.239

Childs, Q. (2022). "This has nothing to do with clouds": A decolonial approach to data centers in the Node Pole. *Commonplace.* https://doi.org/10.21428/6ffd8432.59c985d5

Christian, A. (2020). Turns out dumping data centres in the ocean could be a good idea. *Wired UK*. https://www.wired.co.uk/article/data-centres-underwater-climate-crisis

Derakhshannia, M., Gervet, C., Hajj-Hassan, H., Laurent, A., & Martin, A. (2020). Data lake governance: Towards a systemic and natural ecosystem analogy. *Future Internet*, *12*(8), 126. https://doi.org/10.3390/fi12080126

EROS Data Center Services, (2019) *Data Center Overview.* https://eros.usgs.gov/eros-data-center-services/data-center-overview

Flavelle, C. (2022, June 7). As the Great Salt Lake dries up, Utah faces an 'environmental nuclear bomb.' *The New York Times*. https://www.nytimes.com/2022/06/07/climate/salt-lake-city-climate-disaster.html

Flyverbom, M., & Murray, J. (2018). Datastructuring—Organizing and curating digital traces into action. *Big Data & Society, 5*(2). https://doi.org/10.1177/2053951718799114

Google (n.d). *Cloud storage documentation guide: Landsat data.* Google Cloud. https://cloud.google.com/storage/docs/public-datasets/landsat

Higgins, S. (2008). The DCC curation lifecycle model. *International Journal of Digital Curation*, *3*(1), Article 1. https://doi.org/10.2218/ijdc.v3i1.48

Hogan, M. (2015). Data flows and water woes: The Utah data center. *Big Data & Society, 2*(2). https://doi.org/10.1177/2053951715592429

Johnson, A., & Hogan, M. (2017) Introducing location and dislocation: Global geographies of digital data. *Imaginations 8*(2), 4-7. http://dx.doi.org/10.17742/IMAGE.LD.8.2.1

Jones, N. (2018). How to stop data centres from gobbling up the world's electricity. *Nature, 561*(7722), 163–166. https://doi.org/10.1038/d41586-018-06610-y

Kaplan, J. (2021, February 27). *Jonah Edwards - Internet Archive infrastructure* [Video]. Internet Archive. https://archive.org/details/jonah-edwards-presentation

Kliff, S. (2011, October 20). The Internet is, in fact, a series of tubes. *The Washington Post.* https://www.washingtonpost.com/blogs/wonkblog/post/the-internet-is-in-fact-a-series-of-tubes/2011/09/20/gIQALZwfiK_blog.html

Masanet, E., Shehabi, A., Lei, N., Smith, S., & Koomey, J. (2020). Recalibrating global data center energy-use estimates. *Science, 367*(6481), 984–986. https://doi.org/10.1126/science.aba3758

Miller, H. J. (2010). The data avalanche is here. Shouldn't we be digging? *Journal of Regional Science*, *50*(1), 181–201. https://doi.org/10.1111/j.1467-9787.2009.00641.x

Miller, S. (2021, October 29). Taming the data deluge. *MIT News | Massachusetts Institute of Technology.* https://news.mit.edu/2021/taming-data-deluge-1029

Neef, D. (2015). Digital exhaust: What everyone should know about big data, digitization and digitally driven innovation. *Pearson Education*.

Pendergrass, K. L., Sampson, W., Walsh, T., & Alagna, L. (2019). Toward environmentally sustainable digital preservation. *The American Archivist, 82*(1), 165–206. https://doi.org/10.17723/0360-9081-82.1.165

Reinsel, D., Gantz, J., & Rydning, J. (2018). The digitization of the world from edge to core. *International Data Corporation.*

Roach, J. (2018, June 5). Under the sea, Microsoft tests a datacenter that's quick to deploy, could provide internet connectivity for years. *Microsoft News.* https://news.microsoft.com/source/features/sustainability/project-natick-underwater-datacenter/

Sacasas, L.M. (2021, August 12). The materiality of digital culture. *Comment Magazine.* https://comment.org/the-materiality-of-digital-culture/

Salamone, S. (2003, February 24). Data deluge. *Computerworld*. https://www.computerworld.com/article/2581282/data-deluge.html

Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M., Koomey, J., Masanet, E., Horner, N., Azevedo, I., & Lintner, W. (2016). *United States Data Center Energy Usage Report* (LBNL--1005775, 1372902; p. LBNL--1005775, 1372902). https://doi.org/10.2172/1372902

Siddik, M. A. B., Shehabi, A., & Marston, L. (2021). The environmental footprint of data centers in the United States. *Environmental Research Letters*, *16*(6), 064017. https://doi.org/10.1088/1748-9326/abfba1

Solon, O. (2021, June 19). Drought-stricken communities push back against data centers. *NBC News.*

https://www.nbcnews.com/tech/internet/drought-stricken-communities-push-back-against-data-centers-n1271344

Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, *43*(3), 377–391. https://doi.org/10.1177/00027649921955326

Story, J. (2022, January 20). The data avalanche and what to do with it. *AdWeek.* https://www.adweek.com/programmatic/the-data-avalanche/

Sutherland, T., & Bopp, G. (2023). The Pacific futures of subsea data centers. *New Media & Society, 25*(2) https://doi.org/10.1177/14614448221149944

Thylstrup, N. B. (2019). Data out of place: Toxic traces and the politics of recycling. *Big Data & Society, 6*(2). https://doi.org/10.1177/2053951719875479

UNEP Copenhagen Climate Centre (2020). *Zero carbon emissions in digitalization*. https://www.youtube.com/watch?v=BtHdYciHwns

UNESCO, (n.d.). *Sustainable development and living heritage.* Unesco Intangible Cultural Heritage. https://ich.unesco.org/en/sustainable-development-and-living-heritage

USGS (2011). *About the USGS Water Data for the Nation site.* National Water Information System. https://help.waterdata.usgs.gov/faq/about-the-usgs-water-data-for-the-nation-site

van Bussel, G.-J., Smit, N., & van de Pas, J. (2015). Digital archiving, green IT and environment. Deleting data to manage critical effects of the data deluge. *Electronic Journal of Information Systems Evaluation, 18*(2), Article 2. https://academic-publishing.org/index.php/ejise/article/view/181