https://doi.org/10.2352/issn.2168-3204.2020.1.0.49

Waste Not, Want Not: assessing the environmental sustainability of the University of Houston's digital preservation program

Bethany Scott, University of Houston Libraries, and Diana Dulek, Houston Public Library; Houston, Texas, USA

Abstract

The University of Houston Libraries previously had no data surrounding the environmental sustainability of its digital preservation program. We set out to gather this data and package it in a way that can be communicated easily to stakeholders such as Libraries administration. Additionally, we explore ways that the digital preservation program could become more environmentally sustainable in the future, and we provide actionable recommendations that other digital preservationists can quickly and easily implement to reduce the carbon footprint of their organization's digital preservation program.

Introduction

Although the University of Houston Libraries (UHL) has taken steps over the last several years to initiate and grow an effective digital preservation program – by critically evaluating existing preservation practices, implementing Archivematica for managing and storing preservation packages, and documenting roles and responsibilities for preservation activities through a formal digital preservation policy – until recently we had not yet considered the long-term sustainability of our digital preservation program from an environmental standpoint. In order to address the environmental sustainability of the program, we began gathering information on the technology infrastructure and its energy expenditures in collaboration with colleagues from UHL Library Technology Services and the UH Office of Sustainability. We also reviewed and evaluated the requirements of UHL's digital preservation policy to identify areas where the overall sustainability of the program may be improved in the future by modifying current practices.

Literature review

University of Houston Libraries previously had no data surrounding the environmental sustainability of its digital preservation program. It is not alone, among university libraries, in lacking information on environmental sustainability of digital assets and programs; traditionally, collecting data surrounding environmental sustainability has not been a priority for cultural heritage institutions. Overwhelmingly, literature on the sustainability of cultural heritage organizations' digital programs has been focused on financial and staffing sustainability – how to make the most of limited funding, equipment, and staff time – rather than on issues of environmental sustainability.

For instance, since 2012, the Digital POWRR project has aimed to address the longevity of digital programs by focusing on "preserving digital objects with restricted resources" [1], specifically targeting small- and medium-sized institutions with limited budgets and offering financial assistance to train librarians and archivists in digital preservation basics. Similarly, staffing survey reports released by the National Digital Stewardship Alliance (NDSA) in 2013 [2] and 2017 [3] implicitly correlate an institution's level of preservation staffing with the long-term sustainability of its digital preservation program. And while these administrative and organizational factors are clearly important to sustaining a digital preservation program long-term, reducing the size of our archives' carbon footprint (and first taking the time to calculate what it might be) will arguably become an even more important factor in coming decades.

The existing cultural heritage literature specifically on environmental sustainability concerns can be divided into two distinct fields: research regarding the sustainability of the physical library and archives facilities, and a small, emerging section of research regarding digital preservation in particular.

Sustainability of library facilities and processes

The literature regarding physical archives and their facilities, while not necessarily concerned specifically with digital preservation, provides insights to the methodologies used in assessing environmental sustainability. Additionally, the recommendations provided lay a foundation that can, in some cases, be incorporated into digital preservation practices.

Facilities and the built environment are a major area of consideration when it comes to a library or archives' total energy consumption. A number of articles focus on Leadership in Energy and Environmental Design (LEED) certification, the de facto U.S. standard for green construction. Criteria for evaluating a building's LEED rating include site location, water conservation, energy efficiency, construction materials, and indoor air quality. In a 2003 Library Journal article, "The New Green Standard," Brown highlights examples from the field of libraries; at that time, libraries accounted for 16% of all LEED construction projects [4]. Several follow-up articles in 2007 [5], 2008 [6], and 2009 [7] covered green construction and retrofitting in library buildings.

In the 2010 article "Greening the Library: Collection Development Decisions," Connell applies strategies for reducing a library's carbon footprint beyond library buildings and into the arena of librarians' roles and responsibilities – specifically collection development activities [8]. Connell considers three facets of green collection development: selecting materials that educate on and promote the topic of environmental sustainability, ensuring that weeded materials and equipment are reused or recycled, and critically evaluating the environmental impacts of print vs. electronic resources as a factor in their selection decisions. Because digital collections require electricity each time they are accessed, their carbon footprint is substantial and growing (compared with print materials whose carbon footprint is primarily generated by the paper industry at a single point in time, the time of production). As a result, Connell urges librarians to educate themselves about the environmental impacts of library collections and processes, "mindfully and consistently adopting energy-saving and resource-recycling policies and behaviors."

More recently, in the 2016 article "Archival adaptation to climate change," Tansey addresses environmental sustainability factors related to both the facilities and the activities of the archival profession, arguing that the appraisal, processing, preservation, and long-term planning decisions that archivists make on a day-to-day basis can help to "reduce institutional contributions to climate change and to protect their repositories from climate-change threats" [9].

Tansey provides the definition of sustainability that we use in this paper: "Development that meets the needs of the present without compromising the ability of future generations to meet their own needs." As an example of the lack of consideration given to long-term future needs when making archival decisions for the present, she cites the More Product, Less Process (MPLP) style of archival processing, in which archivists forgo conservation actions (such as rehousing in acid-free folders or item-level handling to stabilize fragile items) in order to manage ever-increasing backlogs, and instead rely on extensive climate controls to passively perform preservation functions – climate control systems that can be highly resource-intensive, contributing to a much greater carbon footprint for the institution.

Taking this definition of sustainability in the context of this paper and its investigation of the UHL digital preservation program's carbon footprint, we are attempting to balance present and future needs by first collecting data with which to make informed decisions, and then by looking critically at digital preservation best practices established over the last two decades and attempting to reconcile those best practices with modern technology available to us.

Sustainability of digital preservation programs

A comprehensive article specifically on the environmental sustainability of digital preservation programs was published in the Spring/Summer 2019 issue of the American Archivist. In the article titled "Toward Environmentally Sustainable Digital Preservation" [10], authors Pendergrass et al. highlight the trend in the cultural heritage field of equating "sustainability" with economic factors such as staffing and funding. In their analysis of the literature that does relate specifically to environmental sustainability, they distinguish three themes: the adaptation of cultural heritage organizations to a changing climate, mitigating the negative environmental impacts of the built environment, and finally, environmentally sustainable strategies for digital practice – a topic that has just emerged in the last 5-7 years and which still requires further interdisciplinary research and development.

The authors go on to explore the full environmental impact of the information and communication technology on which digital archives depend. Extensive support infrastructures are necessary to create usable networked and distributed storage, and discovery and delivery services increase the use of technology by increasing users' online or digital access to collections. While many assessments of the use of computing in libraries and archives focus solely on the electricity usage of computing components, a life-cycle assessment includes additional impacts such as "raw material extraction and refining, shipping at multiple points, manufacture, electricity and cooling during use, and, finally, disposal" – revealing that environmental impacts are in reality much greater than power draw by IT components.

The authors also outline several areas in which changes to digital preservation policies and preservation planning activities can help reduce their power draw and carbon footprint, thereby increasing the sustainability of their digital preservation programs. Strategies for reducing digital preservation's environmental impact through technology include increasing energy efficiency by implementing new technology or modifying existing systems to be less resource-intensive, scheduling high-energy or high-bandwidth tasks for off-peak times of the day or year, and shifting to clean sources of energy when possible.

These are just stopgap measures, though – the authors also call for a paradigm shift in the field of digital preservation as a whole, emphasizing that exponentially growing digital collections may only be sustained long-term through critical reevaluation of the concepts of appraisal, permanence, and availability.

Environmental scan of universities' current sustainability practices

To provide context to our investigation of UH's electricity consumption and sustainability efforts – and how they compare to other universities – we conducted a brief environmental scan of higher ed institutions' current sustainability practices.

Nearly 1,000 institutions have registered with the Association for the Advancement of Sustainability in Higher Education (AASHE) and its Sustainability Tracking, Assessment, and Rating System (STARS). Of those institutions, only five have currently achieved a Platinum rating, the highest rating possible. Institutions are rated against a checklist [11], earning points for sustainable actions in different areas across their campuses.

The STARS checklist does not specifically make mention of campus library buildings or any other particular facility. One specific often library-supported activity that allows institutions to earn or lose points in the STARS checklist is facilitating open access publishing. AASHE is concerned with university sustainability in a holistic sense, thus asking institutions to assess the longevity of their research.

However, the availability of clean and renewable sources of energy on campus is a major criterion in the STARS checklist, and this applies campus-wide (including to library buildings). In the category of "operations," AASHE identifies several activities through which universities can gain detailed information on their carbon emissions and reliance on fossil fuels – and thus can begin to work toward reducing these in the future. For instance, to achieve full marks, universities must have completed an inventory to quantify the institution's Scope 1 and Scope 2 greenhouse gas (GHG) emissions; this includes both direct emissions from owned or controlled sources of energy (Scope 1), and indirect emissions from the generation of purchased energy (Scope 2) [12]. They must also have data on their energy mix, including grid-purchased electricity and electricity from on-site renewables, and they must actively support the development and use of clean and renewable energy sources.

Although only five universities have reached Platinum status, dozens are at the next-highest level, Gold, and are seeking to advance. The University of Houston (UH) holds a Gold ranking and has published a Path to Platinum Plan [13] in which it sets an improvement plan for the university and outlines its sustainability goals. UH's stated goals include a 10% total campus waste reduction, a 15% campus water use reduction, and a 35% reduction in energy expenditures for campus buildings. In the area of sustainability planning, a published climate action plan and a shift to 30% renewable energy on campus are two additional goals for improvement – however, as the report notes, to achieve these goals the Office of Sustainability requires institutional support from the highest level of UH administration, and a culture of sustainability must be adopted by students, faculty, and staff.

In addition to reporting with AASHE, many higher education institutions are also taking the step to formally declare support for climate actions set forth in the Paris Agreement. 353 institutions of higher education have signed the "We Are Still In" [14] declaration of support for climate action. And more than 450 college and university presidents have signed onto the Presidents' Climate Leadership Commitments, agreeing to "work towards carbon neutrality, build resilience, and get students involved in creating climate solutions." [15] To date, seven American universities have achieved carbon neutral status; a 2018-19 impact report from NGO Second Nature showed an additional thirteen have reduced carbon by at least 50%, and fourteen institutions cover 100% of their electricity consumption with renewable energy. [16]

Carbon footprint analysis

The major focus of this research is estimating the carbon footprint of UHL's digital preservation program. The analysis includes three components: 1) conducting an inventory of the equipment used for digital preservation activities; 2) calculating the electricity usage and the sources of energy for local equipment and processes; and 3) investigating offsite storage providers' carbon footprint and any sustainability practices or policies they may have in place.

Inventory of equipment

Utilizing a hyperconverged infrastructure in which compute, storage, and storage networking components are centralized, UHL's digital access and preservation environment is almost 100% virtualized, with all of the major servers and systems for digital preservation – notably, the Archivematica processing location and storage service – running as virtual machines (VMs). The virtual environment runs on VMware ESXi and consists of five physical host servers that are part of a VMware vSAN cluster, which aggregates the disks across all five host servers into a single storage datastore.

VMs where Archivematica's OS and application data reside may have their virtual disk data spread across multiple hosts at any given time. Therefore, exact resource use for digital preservation processes running via Archivematica is difficult to distinguish or pinpoint from other VM systems and processes, including UHL's digital access systems. In this analysis, we are taking a generalized or blanket approach and calculating the power use for all five hosts. This calculation thus represents the energy expenditure for not only the digital preservation system and storage, but also for the A/V Repository and Digital Collections access systems. At UHL, digital access and preservation are strongly linked components of a single large ecosystem, so the decision to look at the overall energy expenditure makes sense from an ecosystem perspective.

In addition to the VM infrastructure described above, all user and project data is housed in the UHL storage environment. The storage environment includes both local shared network drive storage for digitized and born-digital assets in production, and additional shares that are not accessible to content producers or other end users, where data is processed and stored to be later served up by the preservation and access systems. Specifically, with the Archivematica workflow, preservation assets are processed through a series of automated preservation actions including virus scanning, file format characterization, fixity checking, and so on, and are then transferred and ingested to secure preservation storage.

UHL's storage environment consists of two servers: a production unit and a replication unit. Archivematica's processing shares are not replicated, but the end storage share (Archivematica's "storage service" application) is replicated. Again, for purposes of simplification, we generalize that both resources are being used as part of the digital preservation program when analyzing power use. Finally, within UHL's server room there is a pair of redundant network switches that tie all the virtual and storage components together.

The specific hardware components that make up the digital access and preservation infrastructure described above include:

- One (1) production storage unit: iXsystems True NAS M40 HA (Intel Xeon Silver 4114 CPU @ 2.2 Ghz and 128 GB RAM)
- One (1) replication storage unit: iXsystems FreeNAS IXC-4224 P-IXN (Intel Xeon CPU E5-2630 v4 @ 2.2 Ghz and 128 GB RAM)
- Two (2) disk expansion shelves: iXsystems ES60
- Five (5) VMware ESXi hosts: Dell PowerEdge R630 (Intel Xeon CPU E5-2640 v4 @ 2.4 Ghz and 192 GB RAM)
- Two (2) network switches: HPE Aruba 3810M 16SFP+ 2-slot

Electricity usage

Each of the hardware components listed above has two power supplies. However, the power draw is not always running at the maximum available for those power supplies and is dependent on current workloads, how many disks are in the units, and so on. Therefore, the power being drawn can be quantified but will vary over time. With the unexpected closure of the campus due to COVID-19, in conducting this analysis remotely, we compare the estimated maximum power draw based on the technical specifications for the hardware components, the draw when idle, and several partial power draw scenarios, with the understanding that the actual numbers will likely fall somewhere in this range.

Table 1: Estimated power use and greenhouse gas emissions

	Daily Usage Total (Watts)	Annual Total (kWh)	Annual GHG (lbs)
Max	9,094	79,663.44	124,175.71
95%	8,639.3	75,680.268	117,966.92
90%	8,184.6	71,697.096	111,758.14
85%	7,729.9	67,713.924	105,549.35
80%	7,275.2	63,730.752	99,340.565
Idle	5,365.46	47,001.43	73,263.666

The estimated maximum annual greenhouse gas emissions derived from power use for the digital access and preservation hardware is approximately 56.3 metric tons. To put this in perspective, it's equivalent to the GHG emissions from nearly 140,000 miles driven by an average passenger vehicle, and to the carbon dioxide emissions from 62,063 pounds of coal burned or 130 barrels of oil consumed. These figures will serve as an entry point to discussions on the importance of environmental sustainability actions – and our plans to reduce our consumption – with Libraries administration, colleagues in the Office of Sustainability, and other campus leaders.

Offsite storage research

A foundational tenet of digital preservation is the concept that “lots of copies keeps stuff safe” [17] – a concept that is exemplified in NDSA’s Levels of Digital Preservation, in which criteria for a mature, fully realized digital preservation program that sustains its content long-term include that requirement to maintain “at least three copies in different geographic locations, each with a different disaster threat” [18]. To fulfill such requirements, digital preservation practitioners are increasingly relying on offsite storage and “cloud” providers to back up and store redundant copies of their content in case of data loss or other failure of local storage. The energy expenditure to store content with an offsite storage provider, then, becomes another component to consider as part of the total carbon footprint of an institution’s digital preservation program.

The top three commercial providers of cloud storage are Amazon Web Services (AWS), Google Cloud, and Microsoft Azure. In a December 2019 article, Daniel Oberhaus assessed the “relative greenness” of each of these, investigating metrics such as the efficiency of data center infrastructure (lighting, cooling, etc.), the efficiency of servers, and sources of electricity [19]. He describes that all of these providers have already addressed many inefficiencies in the hardware, software, and infrastructure elements of their data centers by running virtual machines, installing custom cooling systems, and more. As a result, these providers can state that their offsite storage services are “greener” or more sustainable than their customers’ local storage – something that AWS explicitly claims [20] – but this is irrelevant for digital preservation practitioners who must duplicate content between both local and offsite storage, and (in the field of higher education) whose institutions are often moving to carbon neutral sources of electricity for their campuses regardless.

Microsoft and Google are both seeking to decarbonize their data centers and offsite storage services through strategies such as increasing their renewable energy portfolios and purchasing renewable energy credits (RECs), which fund the generation of renewable energy in order to offset energy expenditures that are powered by fossil fuels [21]. As an entire company, Microsoft (and not just its Azure cloud) has been carbon neutral since 2012, and even without RECs its data centers run on 60% renewable electricity, which the company plans to boost to 70% by 2023. Microsoft is also making large investments in clean energy projects including hydro and wind power purchase agreements, and in research intended to increase the energy efficiency of current and future data centers. Similarly, Google is increasing its portfolio of clean energy through power purchase agreements, which fund the construction of new renewable energy projects with utilities around the world, expanding the overall availability of renewable resources on the grid. Google Cloud also demonstrates transparency in its global annual carbon accounting, and it is seeking to source carbon-neutral electricity on a “24/7 basis” (in other words, even when the sun isn’t shining or the wind isn’t blowing) [22].

Amazon’s AWS is by far the largest cloud computing provider with well over one-third of the market. Amazon’s most recent sustainability report was released in 2018; at that time AWS exceeded 50% renewable energy usage for its data centers with RECs factored in, and the company made a commitment to reach 100% renewable energy and net zero carbon emissions by 2040. Amazon has renewable energy projects in the United States, Ireland, and Sweden, but the demand for electricity to power its data centers, caused by the expansion of the AWS storage service, greatly outpaced its supply of renewable energy, particularly in Virginia where the core of AWS’s global infrastructure is located. In a 2019 report, Greenpeace details that Amazon’s power demands in Virginia increased by 60% between 2017 and 2019 without adding any new renewable energy, and as a result, its data centers in Virginia are only 12 percent renewable-powered [23].

Although cultural heritage institutions represent a small part of Amazon’s overall customer base, many of these institutions are turning to AWS for offsite storage service; in fact, AWS has its own Public Sector team that is geared toward serving government, education, and nonprofit organizations, and which conducts outreach to partner organizations in the cultural heritage sector such as DuraCloud and Preservica.

In a presentation given at the AWS Preservation Archival Summit in May 2019 [24], Mike Davis of AWS Storage Business Development appeals directly to digital preservation practitioners in outlining reasons they may want to choose AWS for their offsite storage and backup needs. From a cost perspective, AWS’s Glacier Deep Archive storage tier is about $1 per TB per month, pricing which is on par with LTO tape vaulting services that higher ed libraries and archives have traditionally used to back up large sets of data. Higher tiers (in both cost and data availability) in the S3 service “provide periodic data integrity verification and correction,” allowing the service to offer extremely high confidence in the durability of digital assets stored. By relying on AWS’s proprietary approach to data integrity involving self-healing technical infrastructure and cryptographic hash functions crawling its object stores in the background, digital preservationists may choose to reduce the interval between resource-intensive hands-on fixity checks – although the lack of transparency behind the proprietary technology could be a concern for preservation practitioners who wish to understand how and why an asset’s integrity and durability are assured. Finally, AWS offers additional value-add features for cultural heritage organizations, including the availability of AI and machine learning tools and services to automatically enhance metadata on stored objects, run speech-to-text transcription and translation for audio and video assets, and more.

In addition to commercial providers, offsite storage service is also available from cultural heritage organizations. Two providers are Chronopolis [25], hosted at UC San Diego (UCSD), and MetaArchive Cooperative [26], founded by six research libraries and now hosted by Educopia Institute. Both of these organizations specifically brand themselves as digital preservation networks, created and hosted by and for digital preservationists in the cultural heritage sector. As an example, at Chronopolis, data in the network are replicated among partner sites at the National Center for Atmospheric Research, the University of Maryland Institute for Advanced Computer Studies, the Texas Digital Library, and UCSD. At least one of those sites is presently moving towards carbon neutrality: in its 2019 Annual Sustainability Report, the University of California system (of which UCSD is a part) affirmed its goal to be climate neutral and powered 100% by clean energy by 2025 [27].

Since these organizations were founded and built based on concepts and best practices in the field of digital preservation, these networks feature additional services for digital preservationists beyond what commercial offsite storage can provide, including built-in replication and geographic distribution, tracking of preservation metadata, availability of curatorial audit reporting, and transparency of technological infrastructure and community-based standards.

Conclusions

While the findings of our carbon footprint analysis are predicated on our institutional context and practices, and therefore may be difficult to directly extrapolate to other organizations' preservation programs, there are several actionable steps and recommendations that sustainability-minded digital preservationists can implement right away. Getting in touch with any campus sustainability officers and investigating environmental sustainability efforts currently underway can provide enlightening information – for instance, you may discover that a portion of the campus energy grid is already renewable-powered, or that your institution is purchasing RECs.

There are other methods by which digital preservation practitioners can reduce their power draw and carbon footprint, thereby increasing the sustainability of their digital preservation programs. At UHL, existing digital preservation policy documentation outlines file formats and specifications for preservation-quality archival masters for images, audio, and video files that are created through our digitization unit. However, as UHL conducts a greater number of mass digitization projects – and accumulates an ever larger number of high-resolution archival master files – we may want to implement a tiered approach to file format selection, through which we match the file formats and resolution of files created to the scale and scope of the project, the informational and research value of the content, the discovery and access needs of end users, and so on. By choosing to create lower-resolution files for some projects, we would reduce the amount of storage space needed for our digital collections, thereby reducing our carbon footprint.

For instance, we may choose to retain large, high-resolution archival TIFFs for each page image of a medieval manuscript book, because researchers study minute details in the paper quality, ink and decoration, and the scribe's lettering and handwriting. By contrast, a digitized UH thesis or dissertation from the mid-20th century could be stored long-term as one relatively small PDF, since the informational value of its contents (and not its physical characteristics) is what we are really trying to preserve. Similarly, we are in discussions to begin providing an entire archival folder as a single PDF in our access system. Although the initial goal of this initiative was to make a larger amount of archival material quickly available online for patrons, the much smaller amount of storage needed to store one PDF vs. dozens or hundreds of high-res TIFF masters would also have a positive impact on the sustainability of the digital preservation and access systems.

UHL's digital preservation policy also includes requirements for monthly fixity checking of a random sample of preservation packages stored in Archivematica, with a full fixity check of all packages to be conducted every three years during an audit of the overall digital preservation program. Frequent fixity checking is computationally intensive, though, and adds to the total energy expenditure of an institution's digital preservation program. But in UHL's local storage infrastructure, storage units run on the ZFS filesystem [28], which includes self-healing features such as internal checksum checks each time a read/write action is performed. This storage infrastructure was put in place in 2019, but we have not yet updated our policies and procedures for fixity checking to reflect the improved baseline durability of assets in storage. Through considered analysis matching the frequency of fixity checking to the features of the storage media, we may come to the conclusion that less frequent hands-on fixity checks, on a smaller random sample of packages, is sufficient moving forward.

Regarding the environmental sustainability of offsite storage, freely available information and tools allow digital preservationists to assess the carbon footprint for their content stored in AWS. In its sustainability reporting, Amazon states that the following AWS regions are carbon-neutral: US-West (Oregon), Europe (Frankfurt), Europe (Ireland), GovCloud (US-West), and Canada (Central). Sustainability-minded preservationists may wish to choose one of these regions when setting up new backups for their local storage, or if geographic distribution constraints allow, they could explore moving their AWS storage services from a more carbon-intensive region (such as US-East) to a carbon-neutral one.

For a more detailed breakdown, the Green Cost Explorer [29], released in November 2019 by location data platform Mapbox, combines Amazon's public information on which regions and data centers run on renewable energy, with the total carbon use across a user's AWS account provided by the AWS Cost Explorer API. The Green Cost Explorer outputs a breakdown of the dollar amounts and percentages of green vs. grey (or carbon-neutral vs. not carbon-neutral) costs for an AWS account. Using this information, digital preservation practitioners can identify which of their AWS services are contributing to the majority of their grey costs, and potentially transfer some of their workloads or storage to a different region to maximize their use of AWS's sustainable regions.

Future work

When we are able to return to campus, we plan to further refine our electricity consumption research, so that we have exact figures to work with rather than estimates. We would also like to investigate whether changes in preservation processes, such as the reduced hands-on fixity strategy outlined above, can have a positive impact on our energy expenditure – and whether this strategy can still provide a high level of integrity and durability for our digital assets over time. Finally, we would like to take a deeper look at sustainability factors beyond energy expenditure, such as current practices for recycling e-waste on campus or a possible future life-cycle assessment for our hardware infrastructure.

Also, a current goal of UHL's digital preservation program is to begin syncing local Archivematica preservation storage with Amazon Glacier cloud storage in 2020. As part of the process of uploading packages to Glacier, a SHA-256 tree hash must be calculated locally and sent with the package to confirm the integrity and completeness of the uploaded package in Glacier [30]. We plan to capture the pre-upload checksum along with some additional technical and administrative metadata into a local database. From this data, in the future for any given package we will be able to compare the checksum at time of upload, the current checksum of local data (in both the production storage and replication storage units), and the checksum retrieved from the Glacier vault inventory for the package stored in Glacier (which is updated at some unstated interval over time). The availability of this full checksum verification as a baseline, as well as local and Glacier fixity reporting options, will also factor into policy revisions regarding the frequency of future fixity checks. And, of course, we plan to strongly recommend that a carbon-neutral region is selected for our AWS Glacier storage.

References

[1] A. Rinehart, P. Prud'homme, and A. Huot, "Overwhelmed to action: digital preservation challenges at the under-resourced institution," OCLC Systems & Services Journal, vol. 30, no. 1, pp. 28-42. (2014).

[2] W. Atkins et al., “Staffing for Effective Digital Preservation 2013: An NDSA Report,” National Digital Stewardship Alliance. Dec. 2013.

[3] W. Atkins et al., “Staffing for Effective Digital Preservation 2017: An NDSA Report,” NDSA. Feb. 2018. Available: https://doi.org/10.17605/OSF.IO/3RCQK [Accessed Apr. 15, 2020].

[4] B. Brown, “The new green standard: with the LEED rating system in place it is easier to make sure your new library saves money as it treads lightly on natural resources,” *Library Journal*, vol. 128, no. 20, pp. 61-65. 2003.

[5] J. Mikkelson, “Going for the gold: Building a sustainable LEED library,” *OLA Quarterly*, vol. 13, no. 4, pp.25-28. 2007.

[6] R. Kuzyk, “Going green: LJ’s second design institute tackles the ins and outs of sustainable buildings,” *Library Journal*, vol. 133, no. 9. 2008.

[7] P. Gisolfi, “A sustainable library, inside and out,” *Library Journal: Library by Design*, vol. 8. 2009.

[8] V. Connell, “Greening the Library: Collection Development Decisions,” *Endnotes: Journal of the New Members Round Table*, vol. 1, no. 1. May 2010.

[9] E. Tansey, “Archival adaptation to climate change,” *Sustainability: Science, Practice, & Policy*, vol. 11, no. 2. 2015.

[10] K. Pendergrass et al., “Toward Environmentally Sustainable Digital Preservation,” The American Archivist, Vol. 82 No. 1, pp. 165-206. 2019.

[11] AASHE, “STARS 2.2 Credit Checklist.” Available: https://stars.aashe.org/resources-support/technical-manual/ [Accessed Apr. 15, 2020].

[12] GHG Protocol, “Scope 2 Guidance.” Available: https://ghgprotocol.org/scope_2_guidance [Accessed Apr. 15, 2020].

[13] UH Office of Sustainability, “University of Houston Path to Platinum Plan,” Fall 2018.

[14] “We Are Still In” Declaration. Available: https://www.wearestillin.com/we-are-still-declaration [Accessed Apr. 15, 2020].

[15] M. Earls, “Colleges commit to carbon neutrality. Getting there is hard,” E&E News, Oct. 3, 2019. Available: https://www.eenews.net/stories/1061192375 [Accessed Apr. 15, 2020].

[16] Second Nature, “2018-2019 Impact Report.” Available: https://secondnature.org/wp-content/uploads/2018-2019-Second-Nature-Impact-Report_FINAL.pdf [Accessed Apr. 15, 2020].

[17] Stanford University, “LOCKSS History.” Available: https://www.lockss.org/about/history [Accessed Apr. 15, 2020].

[18] NDSA et al., “2019 Levels of Digital Preservation.” Available: https://doi.org/10.17605/OSF.IO/QGZ98 [Accessed Apr. 15, 2020].

[19] D. Oberhaus, “Amazon, Google, Microsoft: Here’s Who Has the Greenest Cloud,” *Wired*, Dec. 10, 2019. Available: https://www.wired.com/story/amazon-google-microsoft-green-clouds-and-hyperscale-data-centers/ [Accessed Apr. 15, 2020].

[20] N. Sahlstrom, “Reducing carbon by moving to AWS,” Nov. 26, 2019. Available: https://blog.aboutamazon.com/sustainability/reducing-carbon-by-moving-to-aws [Accessed Apr. 15, 2020].

[21] EnergySage, “Renewable Energy Credits,” Jan. 10, 2020. Available: https://www.energysage.com/other-clean-options/renewable-energy-credits-recs/ [Accessed Apr. 15, 2020].

[22] M. Terrell, “The Internet is 24x7. Carbon-free energy should be too,” Oct. 10, 2018. Available: https://www.blog.google/outreach-initiatives/sustainability/internet-24x7-carbon-free-energy-should-be-too/ [Accessed Apr. 15, 2020].

[23] Greenpeace, “Clicking Clean Virginia: The Dirty Energy Powering Data Center Alley,” Feb. 13, 2019. Available: https://www.greenpeace.org/usa/reports/click-clean-virginia/ [Accessed Apr. 15, 2020].

[24] M. Davis, “Persisting Digital Assets for Durability and Value,” May 8, 2019. Available: https://zoom.us/recording/play/-cSv2JD5itE1qLu8Pb6fTdg5VT4xuV8BIaqrcclJuSHlEMoywr2Uetm9ysyz6QUA [Accessed Apr. 15, 2020].

[25] Chronopolis, “About, Mission Statement, Services, TRAC Certification.” Available: http://libraries.ucsd.edu/chronopolis/about/index.html [Accessed Apr. 15, 2020].

[26] MetaArchive Cooperative, “About Us.” Available: https://metaarchive.org/about/ [Accessed Apr. 15, 2020].

[27] University of California, “Annual Report on Sustainable Practices,” 2019. Available: https://sustainability.ucsd.edu/_files/UCAnnualSustainabilityReport.pdf [Accessed Apr. 15, 2020].

[28] FreeNAS, “ZFS Primer.” Available: https://www.ixsystems.com/documentation/freenas/11.3-U2/zfsprimer.html [Accessed Apr. 15, 2020].

[29] “Green Cost Explorer – climate related spend analysis for AWS.” Available: https://github.com/thegreenwebfoundation/green-cost-explorer [Accessed Apr. 15, 2020].

[30] “Computing Checksums – Amazon Glacier Developer Guide.” Available: https://docs.aws.amazon.com/amazonglacier/latest/dev/checksum-calculations.html [Accessed Apr. 15, 2020].

Author Biography

Bethany Scott is Digital Projects Coordinator at the University of Houston Libraries, where she is a contributor to the development of the BCDAMS ecosystem incorporating Archivematica, ArchivesSpace, Hyrax, and Avalon. As a representative of UH Special Collections, she contributes knowledge on digital preservation, born-digital archives, and archival description to the BCDAMS team. She received an MS in Information Studies from the University of Texas at Austin.

Diana Dulek is a Systems Librarian at Houston Public Library. She will graduate with her MLS from Texas Woman’s University in 2020. She is interested in sustainability in the archives, social justice, and the preservation and discoverability of marginalized voices.