

Emily Nabasny
Digital Preservation
Kara Van Malssen
December 13, 2013

Dark Archives: Deep Sixing in the Digital World

Dark archives, like so many things, are heavily subject to an individual's opinion and understanding. There are many different kinds of organizations, institutions, and businesses that house and control archives and digital collections, and there are many different terms for these places: repository, archive, library, network storage, DAM, etc. But what does it really mean when one of those is "dark", and how will that affect the continuing preservation of the content it holds?

Archives function as long-term storage for records, audiovisual materials, and various other ephemera. Archives also serve as to collect, organize, preserve, and allow access to these materials through a central access point. Dark archives function as a repository for materials and information for long-term preservation, storage, back-up, and even in some cases as a failsafe in the event the material becomes unavailable elsewhere.

Dark archives step beyond simply having some restricted materials in their holdings, and enter the realm of having unprocessed and unknown materials because of copyright, privacy, security, or other such reasons. In many cases this is an unintentional result of a processing back log that keeps materials unknown, unprocessed, undocumented, and therefore inaccessible; a fate that may in fact affect the life of the material. These types of archives or collections are not unknown in the world of physical materials, there are items in nearly every archive that are restricted for one reason or another, however in this increasingly digital age it has become so expected to be able to search and access information easily. That it is sometimes forgotten what it's like to search and search for information and never get it. We live in a world where nearly every aspect of a person's life is shared digitally, every day we converse digitally, and yet we are unable to keep record and search for proof of important historical moments.

Are there records of what is in these dark archives? Will it be released at a later date? Perhaps waiting on a restriction to expire? Do these archives only reveal themselves when material is thought to be lost? How will we know what is there? Dark archives are frustrating for archivists, so many unanswered questions. How do dark archives will fit into the digital archive setting, and how will they function within the demands of digital preservation?

What is a Dark Archive?

Dark archives are a hard entity to define. Generally it is an archive or repository with content that is not accessible to the public. It is simple to say that they are a repository for materials and information that needs to be kept inaccessible for a specific reason, or to say that it is an archive with a large backlog of unprocessed materials keep, but it is not that simple. Yes, they can be places where information we thought lost forever lives on, but they are also places where rights holders and depositors can keep

their materials and know it will survive, a place where materials that are intentionally meant to be kept hidden will be continually managed and undergo preservation needs.

By their own design, dark archives focus on preservation first, and are next to impossible to see. They are only truly noticed and seen when they are release content, and in doing so often disrupt the perception of history. Unveiling dark secrets and hidden truths; allowing information to spread out into the world. While that makes it seem far more glamorous that it is, it is not far off from the reality that comes with the revelation of a dark archive. They are unlikely to reveal unknown scientific truths or scrolls from Alexandria, but that is why the documentation of content is extremely important in this now digital world we find ourselves in. A backlog of unprocessed material in the in the digital world is bad. With digital there are no tangible objects to look at and see. You cannot turn page or use a microscope for a closer look at content. Processing these materials as soon as possible is a must when there are no physical backups or copies to pack away in a vault or spread throughout the world. It is all files names and bits, putting our faith in the code of the digital world.

There are three major requirements of digital preservation: bit preservation, accessibility and usability, and sustainability. These three elements are crucial to the long-term preservation and survival of digital content, because without them the content risks data corruption and permanent loss. Bit preservation – done through checksums – works to ensure that all of the bits – 1’s and 0’s - in the data stream are intact and correct, and have not experienced bit corruption, as even one flipped byte can cause a file to become unreadable or to loose data. Maintaining accessibility and usability of data is important even in a dark archive when content is not being accessed outside of the archive because those within the archive need to be able to access content in order to complete checksums, send reports to the contributors, and so that if these archived copies are ever needed in the future that they will be accessible through the newer formats and software. Finally, sustainability is an extremely important factor that – while still debatably defined – asks for consideration of the cost benefit, content value, and physical analysis of the ability to preserve content long-term. Cost, time to preserve, maintenance, and archival stability are just some of the pieces that fall under the larger umbrella of sustainability, especially in terms of a dark archive where getting consistent funding or support could prove difficult.

Dark archives are also accompanied with the major question of “dark to whom?” Who is it really that wants access to these materials, and is unable to reach them. Typically it is the general public, the people around the world, who are unable to access materials. In some cases the rights holders are able to access materials through specific request or a given password, and in some cases not. More often than not though, it is only the archive staff that has access and knowledge of the archive. Keeping in mind also, that “unintentional” dark archives or dim archives are considered dark because of the dark content they hold.

Dark content, a term that is very relevant to the discussion on dark archives, because it is often the result of an archive having dark data – or dark content – that it is considered to be dark. However it is often that archives are considered to be “unacknowledged” dark archives simply due to a portion of their holdings being inaccessible. This separates dark archives from dark content.

While this moniker may typically be raised in regards to scientific or research material, or in the case of copyright, it can be readily applied to materials that are dark for any number of reasons. Particularly when considering that within a dark archive it is not simply content that is dark, but their data on procedures, protocols, technical data, and all metadata as well, that is dark with little expectation of coming to light.

In mid-2013 at a data archiving facility at Columbia University, and scientific publisher Elsevier teamed up to offer an award to the individual who could best exemplify how data-preservation techniques are utilized to uncover new earth sciences discoveries. This competition is meant to showcase not only the innovative ways in which digital information is stores, preserved and accessed, but also to display how the information is used.¹ In this instance it is not a “dark archive” that is hidden away, but rather the data itself that has been made “dark.” Scientists, and scientific studies and research are deterred from accessing data that originated from failed or abandoned projects and experiments that could ultimately be extremely helpful in new studies and uncovering new research projects. This contest challenged participants to describe the data before and after it was “rescued”, how it would be used, how it was rescued, and what models and standards are used.

Aside from its “dark” and inaccessible status, this type of material also makes prime example of a major problem that can occur with archives, not being properly cataloged and archived. The material is independently inaccessible but even if one was granted access, the materials are not cataloged, archived, or accessible.

This is a similar process as to what could happen with every “dark” repository. This information and more is important to the continued operation and care of archives and their content. Even if a specific list detailing what is being stored is unavailable to the public or third parties, there must be records detailing standards, metadata information, and all other technical information needed to operate, save, and migrate these archives, that could be made available if needed.

Portico

Portico is a digital archive that grew out of an initiative started in 2002 by JSTOR as a solution for the preservation of electronic scholarly journals. The initiative transitioned to ITHAKA in 2005, and has since become a sustainable digital archive that serves the academic community, while also allowing publishers and libraries to feel safe as they to rely more and more on digital content. In this effort to continually preserve content for future generations Portico is an established TRAC certified dark archive.² Only those working within Portico have direct access to content, while contributors – libraries and publishers – can check content status through a given login.

Portico’s main goal of addressing the need of a permanent archive of electronic scholarly content means that their archive is continually growing as they preserve more

¹ Martineau, Kim. "Contest Aims to Bring 'Dark Data' into Digital Archives." *State of the Planet: Blogs From the Earth Institute*. April 29, 2013. <http://blogs.ei.columbia.edu/2013/04/29/contest-aims-to-bring-dark-data-into-digital-archives/>.

² Center for Research Libraries. "Center for Research Libraries: Report on Portico Audit Findings." *Center for Research Libraries*. January 10, 2010. <http://www.crl.edu/sites/default/files/attachments/pages/CRL%20Report%20on%20Portico%20Audit%20010.pdf> (accessed 2013).

and more content. Portico uses file normalization and migration for the continuing of digital preservation for the content it houses and, best practices for all acts that take place behind the closed doors. Portico does however; publish their preservation process, metadata recording policies, and intention to pass along their content in the event that they themselves cease to exist. Where many dark archives keep this information secret, those services are one of Portico's major selling points. This information builds confidence in contributors and faith in long-term sustainability and usability. A choice that lends itself to the continuing sustainability of the Portico archive as it is entirely set for the purpose of preservation and preservation action. When a service is fully presented and understood, it is more likely to be supported and invested in, it is sustained

A major design of Portico that makes its approach interesting is that Portico is both a "dark" archive and a "light" archive. Meaning that when content is still available outside of Portico through other sources it cannot be accessed through the archive. The archival content becomes "light" when participants audit the archive, and when content is made available as a result of "trigger event", which signals the content is unavailable anywhere else. This makes researchers pleased that content is preserved and can be made available, and keeps it unavailable until absolutely necessary, per the rights agreement. The initial agreement between publishers and Portico, gives Portico the right to obtain, normalize, migrate, and exhibit content for preservation purposes.

Portico's mission to continually store, migrate, and provide access to these materials gives libraries, publishers, and scholars the confidence to be more comfortable when choosing to rely on digital content. While this is a dark archive, it is an excellent model for how dark archives should operate, intentional plans and work, full processing, and all signs pointing to longevity of content.

As more publishers and researchers become rely on electronic content, the continually growing problem of long-term access becomes ever more relevant. What would happen if a publisher goes out of business? Or if they no longer offer back issues? Is the archive sustainable? Portico is well over 8 years into this project and their digital preservation approach mixing long-term commitment from organizations with long-term content management is working very well. Their main goal of preserving scholarly content is overseen by a Board of Trustees who represent the academic community (university presidents, librarians, publishers, and faculty members), ensures the review and guidance of Portico's policies, and the use of best practices in the continuing digital preservation of the content.

Content of the project includes not only born digital content such as academic e-book, scholarly e-journals, and digitized historical collections (d-collections) as well. Portico continues to preserve over 227,000 thousand e-book titles, 16,000 e-journal titles, and 72 digital collections from 849 libraries and 219 publishers from around the world. (A detailed list of the content in the archive can be found online.³) Even when content is made accessible, once content is deposited in the archive, it remains there.

As the preservation of digital content can present many technological challenges, content is preserved through a migration strategy that focus on 2 key points: the first, to identify preservation metadata at the initial point of the preservation process, and the

³ Portico. *Portico- A Digital Preservation and Electronic Archiving Service*. 2005-2013. <http://www.portico.org/digital-preservation/archive-content-access/access-to-archived-content> (accessed 2013).

second, to only migrate content at the point when it is necessary. This further is detailed in a 5-step preservation approach that is implemented when ever new content is ingested. The 5-steps of preservation are: 1) Preservation Planning, 2) Receipt & Inventory Management, 3) Processing & Archival Deposit, 4) Monitoring & Management, and 5) Content Delivery.

The preservation process automatically begins when Portico receives the source files from the publishers; analyzing the files and putting them into an archival package separate from the original file (step 1: Preservation planning). Operations staff converts the full-text files into archival format based on the Journal Archiving and Interchange DTD created by the National Library of Medicine (NLM). As it is source files, and not delivery files, that the archive takes in, high-quality HTML, PDFs, SGML or XML text files, images, and supplemental data such as video and audio is preserved as well and made available. Preservation plan is done for each individual ingest and are developed based on the content and the needs of the provider and libraries. New tools necessary to the execution of the plan may also be created if they are necessary to: retrieve or load content, repackage the content, or migrate files in certain formats. If files cannot be converted, the files are checked for bit corruption and go through bit preservation. The content is then manually or automatically transferred by way of: portable media, File Transfer Protocol (FTP), Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) or Open Archives Initiative Object Reuse Exchange (OAI-ORE), or through Portico's custom software to Portico, and loaded into the storage area (step 2: Receipt & Inventory Management).

As the content is processed and deposited in the archive (step 3: Processing & Archival Deposit), preservation metadata is created following the OAIS model, and includes all DTDs, schemas, and documentation necessary to understand the content. The metadata also includes: validating files for the formats, checksums, technical and descriptive metadata pulled from the source file, and a record of all actions taken on the content. Express care is taken in each stage to get the best results and archive the most information. Portico preserves the intellectual content of the files, and the resources themselves by using a modified version of METS to track actions taken on the source file. All this information is monitored by Portico and supports the longevity and sustainability of the files.

After ingest, and throughout the life of the files, the operations staff monitors the condition of the content using sample data (step 4: Monitoring & Management). Files are occasionally audited, and during audits content can be replicated, checked for problems, repaired or replaced if corrupted, audited, updated as per preservation needs, or migrated to new formats. Portico's archive is also replicated offline with multiple copies that are geographically separated. The final stage of Portico's preservation plan (step 5: Content Delivery) grants access of preserved content to Portico supporters when such a time arrives that those titles are no longer available from the publisher or another source. This circumstance known as a "trigger event," and is very unique to Portico and a few other systems. Access is also available to libraries when a claim has been filed and okayed by the publisher.

A major design of Portico that makes its approach interesting is that Portico is both a "dark" archive and a "light" archive. Typically dark, the archive becomes "light" when participants audit the archive, and when content is made available as a result of

“trigger event”. (There is a list on Portico’s website of triggered content.⁴) Participating publishers and libraries are given password access to the archived content for the purpose of auditing, verification, and use at any time after ingest. Non-participants can never access Portico’s content. As a result of this design, Portico secures the right to obtain, normalize, migrate, and exhibit content for preservation purposes.

In 2010 the Center for Research Libraries (CRL) certified Portico as a trustworthy digital repository following an audit they conducted in 2009. It is actions such as these that reassure content owners and support the sustainability of the archive. Libraries and publishers around the world know that they can rely on Portico when triggered events lead to loss of access. It is amazing to see so many participants with faith and support for Portico and interest in archiving content. At Drew University, the library was hesitant at first because they were concerned with availability, but now they use Portico and rely on its archive to fulfill their users needs.⁵ Portico is assisting not only in the preservation of content but to the global discussion of archiving as well.

Overall Portico remains an important dark archive model of digital content for the preservation community. Their mission to continually store, migrate, and provide access to these materials gives libraries, publishers, and scholars the confidence to be more comfortable when choosing to rely on digital content and the longevity of the materials. Portico’s storage and archive practices are carefully thought out and executed and the actions taken on content continually reveal the interest in long-term preservation over immediate access.

DAITSS (Dark Archive in the Sunshine State)

Another example of a successful and popular repository that is openly recognized as a “dark archive” is DAITSS.

DAITSS is an OAIS based preservation software application developed by the Florida Center for Library Automation (FCLA) and for use by the Florida Digital Archive (FDA), a long-term preservation repository. This application is designed “specifically to ensure the long-term renderability of authentic digital materials,”⁶ The FDA shares this digital repository with eleven universities in the Florida public university system. (All of which put material with DAITSS that is upon ingest inaccessible.)

DAITSS develops active preservation strategies, maintains standardized metadata, maintains data on digital provenance, and continuously performs fixity checking on multiple stored copies as a means to longevity. This system offers automatic support for submission, ingest, archival storage, access, withdrawal, and repository management. The preservation protocol combines bit-level

⁴ Portico. *Triggered Content*. 2013. <http://www.portico.org/digital-preservation/the-archive-content-access/access-to-archived-content>. (ITHAKA 2005-2013)

⁵ Connors, Linda. *Looking to the Future*. 2008. <http://www.portico.org/digital-preservation/who-participates-in-portico/what-participants-say-about-portico/looking-to-the-future> (accessed 2013).

⁶ Caplan, Priscilla. *DAITSS, an OAIS-based preservation repository*. Gainesville: Florida Center for Automation, 2010. 1.

preservation, format normalization, and forward format migration to keep the files in the preservation process and useable.

Following the model for an OAIS system, the DAITSS system “can accept a Submission Information Package (SIP), transform the SIP into a stored Archival Information Package (AIP), and transform the AIP into a Dissemination Information Package (DIP) on request.”⁷ This is all accomplished through the implementation of 4 of the 6 functions of OAIS: Ingest, Data Management, Archival Storage, and Access. The other 2 elements of OAIS not performed by DAITSS – Administration and Preservation Planning - are performed by FDA staff with the support of the reporting and data management functions in DAITSS.

As of June 2011, the repository held “290,000 packages comprising 39.1 million files and taking 87 TB storage for a single copy.”⁸ FDA estimates that on average four to five terabytes are ingested each month. That is a lot of data to take in and a lot to maintain in a usable, and sustainable state. That is why preservation plan and worked out workflow is necessary.

Like Portico, DAITSS’s first and most central job is ongoing preservation that insures the usability of the files. Their job is not access, which is done through other entities. The level of preservation planning found within organizations such as this is a good standard from which other archives should take note. In these instances, these companies are not offering just the storage that many institutions want but also the preservation and continued preservation work that they truly need.

The participating institutions can access only their own files within DAITSS. Requests can be sent to DAITSS by the institutions for specific content from their own collection, but they cannot access the materials submitted from the other Universities. This design makes the content within the archive dark by the choice of the archive being geared toward preservation – not access – purposes, and by the design that only the giving institution can access their files.

Other Archives

More commercial storage options like Amazon Glacier - and the many cloud options - promote their “safe storage”, data archiving, and back up. These archives, however, are designed for access and storage over preservation though. Glacier in particular is set up as a low-cost storage for back up, and it –in particular – is designed for infrequent access and has long data retrieval times.⁹ All archiving services have initial and ongoing costs; however is it worth it in this case? To pay money and have little access to materials and ultimately not really know if you will actually be able to use them when you want to?

Government archives are great examples of “unacknowledged” dark archives that are not always up to archival and preservation standards. We know that the archives exist and are actively growing, but we are unable to access them and have no idea what they hold. In the case of the Library of Congress, archiving and preservation standards are some of the best but the archive itself makes for an interesting example of a combined library, repository, and

⁷ Florida Digital Archive. *Chapter 1: DAITSS Overview*. October 25, 2011. http://daitss.fcla.edu/sites/daitss.fcla.edu/files/Chapter_1_DAITSS_Overview.pdf.

⁸ Ibid.

⁹ Amazon. *Amazon Glacier*. 2013. <http://aws.amazon.com/glacier/>

archive that functions as a light, dim, and dark archive. The Library of Congress holds materials that can be actively borrowed, ones that can only accessed only by, and materials that cannot be accessed at all by those outside the institution. A majority of the LOC's holdings fall into the dark archive category as it continues to intake fast quantities of materials from the copyright offices, and acts as their eternal repository. This combination of functions very nicely exemplifies what occurs in smaller libraries, and where the dilemma and conflict around both digital and dark archives occurs when dealing with so many materials with different policies and requirements.

Unfortunately not all government "dark" archives are like the LOC and maintain excellent preservation work. Many could learn from the practices upheld by the LOC, Portico, and DAITSS as it is evident in recent news articles over the last few years that archives in both the NSA and NASA have experiences preservation problems. NASA, a few years ago, experienced the unfortunate problems of both have lost unique space footage, and encountering the problem of not being able to view material because the format had become outdated and obsolete.¹⁰ The NSA - while not avidly discussing any issues – I would imagine is being entirely overwhelmed with the amount to of digital data and information that they are ingesting now that the number of satellites and cameras observing the movements of the world have increased so exponentially.

Simply because an item is in an archive does not mean that in 20 years from that moment you will be able to look at it like it was yesterday. Technology changes, and it changes quickly. When it does change, the old technology slowly stops and ultimately begins to disappear, be lost in time. While dark archives keep content safe from prying eyes, it is not a free pass to forget what is there, and continues to build upon it. This content needs the same attention and work that materials in open access archives need.

Sustainability

However inconvenient and removed dark archives can be, there are practical reasons for "deep sixing" materials. Dark archives provide good holding locations for: government files, materials dealing with copyright, confidential documents, anything that is desired to be kept private or unpublished, and more. Consider copyrights and intellectual control. If an archive fails to get some rights from the rights holder then the archive has to "deep six" that material until such a time that they are able to allow access to it. This is in part why Portico is TRAC certified, rights are part of the certification criteria and Portico makes sure to get those rights. Not only so that they can allow access to the material if the time arises, but also so that they can act on any preservation need that the content requires. However even in all of these instances, there remains the growing concern of these repositories and archives being able to properly care for and sustain all of the content that they are taking in.

As with all archives - and especially digital archives – there is a large questions surrounding sustainability. Is the archive worth the money and time it costs? What does it mean to be sustainable? In order to be sustainability costs need to be covered and content/services need to be continually available and accessible. "Despite the great value of the content being entrusted to a digital format, the business models that will ensure

¹⁰Kaufman, Marc. "The Saga Of the Lost Space Tapes." *The Washington Post*. 2007.
<http://www.washingtonpost.com/wp-dyn/content/article/2007/01/30/AR2007013002065.html>.

long-term access to and preservation of this material are still unclear.”¹¹ Can archives like Portico and DAITSS be used as a sort of ideal business model and workflow formula? If the funders, patrons, and even the archives do not know what it has, how can they be sustainable? That all strongly depends on the main goals and functions of the digital archive. For sustainability it is not always what the content is and whether or not that is worth saving, as it is making a case for why the archive itself should survive and be invested in. Archives need to promote themselves and their actions towards preservation and long-term access, not just what collections they keep locked away.

One of the other more problematic issues with digital materials is knowing who has copies of what. Digital materials can so easily be copied, transferred, moved, and stored in a variety of locations, but how do we know that is happening? How do we know that if a tsunami were to hit the main servers of some large corporation, that those documents and reports would be available from another source in another country? Being aware of what is where is a huge part of sustainability unto itself, and in institutions with lots of projects and a changing work flow that can be very difficult. Which is why archives should publicly release a list of their holdings, so that even if something is not accessible, it is known that that item is being preserved.

With the question of sustainability naturally comes the question of worth. When spending the amount of money and time on preserving materials that is spent in archives, it is natural to reevaluate the value and worth of what it is you are saving. There are many elements to factor into determining the value and worth of content that as many of them must be considered as possible. Does it have cultural, historical, or sentimental value? Does it have monetary value? Format obsolescence and developing software forces the reconsideration of the value of digital elements, but it is merely a major - not the only - contributing factor. In the digital world value may not be high money wise, but culturally there could be great value in things and as technology develops, the trick will be to be able to determine if the content or the format is more valuable, especially as storage becomes an issue. Storage was always an issue for physical materials, but now it is for digital as well. As digital material continues to grow exponentially by the second, how do these archives intend to keep up the storage demands? DAITSS reports that it takes in four – five terabytes a month, that is a lot of content and space required, and that number is only going to increase from here.

So with all these considerations what is to be done in terms of sustainability? Institutions and organizations are always keeping their mission and funding in mind. It works best for them to acknowledge funding sources, and possible future funding sources from the initial creation of the archives, because it could be difficult to find the funding later. But when archives only run into sustainability troubles as they grow it may become beneficial to – as mentioned before – promote the preservation work and longevity of the materials, over the materials themselves.

What does this mean for digital preservation?

¹¹ Maron, Nancy L., K. Kirby Smith, and Matthew Loy. "Sustaining Digital Resources: An On-the-Ground View of Projects Today: Ithaka Case Studies in Sustainability." Ithaka, 2009.

The final, culminating question behind all of this is “what does this mean for digital preservation?” There are no clear answers for what affect dark archiving will have on the preservation of digital materials, and there will probably never be. Everything will always be changing and evolving. Working out solutions to problems and then finding new problems.

Already digital content – both born digital and digitized – is being carefully controlled and released in many archives. Living in a world where we interact and learn from the digital world every day, it can be hard to grasp the concept or idea that digital information is unavailable. If it is digital why is it not streaming? Why is it not instantly accessible?

There are positives to these digital dark archives though. Mainly, the well-structured archives are preserving the content they hold so that it could potentially be accessed in the future. There are also several factors concerning dark archiving may cause big archiving problems:

- 1) Lack of preservation planning
- 2) No access – intellectual property rights, legal reasons, privacy, government
- 3) Exponential growth
- 4) Cost
- 5) Speed of hardware/software development

And how these problems are faced and dealt with now will influence future standards and work in the field. It is hard to create and maintain digital preservation standards, as the digital world is continually growing and changing every day, but workflows and business models can be set into place that will work no matter what the software.

Formats are predicted to last 5 to 10 years, but new formats, software, and hardware comes out every year. In terms of digital preservation this means that as archivists and consumers we need to be continually aware of what these storage companies are offering. Storage is not synonymous with access, preservation, or back up, but all of these elements are needed to ensure content survival.

Conclusion

Dark archives are not evil. These archives are not a sign of the oncoming doom of digital preservation, and they are not the end of open access. But that does not mean they should not be carefully watched. Living in a world, in a culture, that relies so heavily on the digital world and immediate access to all things has resulted in two major actions: first, we expect that everything and anything we want is instantly accessible online so we get angry when it is not, and two companies and organizations are hyper aware of the piracy of digital materials and therefore take numerous actions to safeguard their copyrighted material. As a result of these two actions, dark archives could create quite a rift in between patrons and rights holders.

Sustainability, preservation, and access are issues that will forever haunt archives and repositories, as they are the true driving force behind the existence of the preservation and archiving field. We want to preserve so that we can later access. And we do not want to hear that somewhere along the way that the “ball was dropped” what we thought was safe has been lost forever.

Bibliography

- Aryis, Paul. *What Participants Say*. <http://www.portico.org/digital-preservation/who-participates-in-portico/what-participants-say-about-portico> (accessed 2013).
- Amazon. *Amazon Glacier*. 2013. <http://aws.amazon.com/glacier/>
- Caplan, Priscilla. *DAITSS, an OAI-based preservation repository*. Gainesville: Florida Center for Automation, 2010.
- Center for Research Libraries. "Center for Research Libraries: Report on Portico Audit Findings." *Center for Research Libraries*. January 10, 2010. <http://www.crl.edu/sites/default/files/attachments/pages/CRL%20Report%20on%20Portico%20Audit%202010.pdf> (accessed 2013).
- Chris. "Monday, May 06, 2013 Light, Dark and Dim Archives: What are they?" *Digital Preservation Matters*. 2013. <http://preservationmatters.blogspot.com/2013/05/light-dark-and-dim-archives-what-are.html>.
- Connors, Linda. *Looking to the Future*. 2008. <http://www.portico.org/digital-preservation/who-participates-in-portico/what-participants-say-about-portico/looking-to-the-future> (accessed 2013).
- Fenton, Eileen. "Portico: A Collaborative Approach to Preservation." *Digital Preservation Europe*. 2007. http://www.digitalpreservationeurope.eu/publications/portico_briefing_paper_def.pdf (accessed 2013).
- Fitzpatrick, Kathleen. *Planned Obsolescence: Publishing, Technology, and the Future of the Academy*. New York: New York University Press, 2011.
- The Florida Center for Library Automation. *DAITSS Digital Preservation Repository Software*. 2011. <http://daitss.fcla.edu>.
- Florida Digital Archive. *Chapter 1: DAITSS Overview*. October 25, 2011. http://daitss.fcla.edu/sites/daitss.fcla.edu/files/Chapter_1_DAITSS_Overview.pdf.
- Kaufman, Marc. "The Saga Of the Lost Space Tapes." *The Washington Post*. 2007. <http://www.washingtonpost.com/wp-dyn/content/article/2007/01/30/AR2007013002065.html>.
- Kirchhoff, Amy, and Eileen Gifford Fenton. "Archiving Electronic Journals: An Overview of Portico's Approach." *Portico - A Digital Preservation and Electronic Archiving Service*. August 2006. <http://www.portico.org/digital-preservation/wp-content/uploads/2009/12/PapersFromPortico.1.Overview.pdf> (accessed September 19, 2013).

Kuny, Terry. "A Digital Dark Ages? Challenges in the Preservation of Electronic Information." 1997. <http://archive.ifla.org/IV/ifla63/63kuny1.pdf>

Leaders. "Bit Rot." *The Economist*. 2012. <http://www.economist.com/node/21553445>.

The Library of Congress. *Audio-Visual Conservation at the Library of Congress*. 2012. <http://www.loc.gov/avconservation/>.

Maron, Nancy L., K. Kirby Smith, and Matthew Loy. "Sustaining Digital Resources: An On-the-Ground View of Projects Today: Ithaka Case Studies in Sustainability." Ithaka, 2009.

Martineau, Kim. "Contest Aims to Bring 'Dark Data' into Digital Archives." *State of the Planet: Blogs From the Earth Institute*. April 29, 2013. <http://blogs.ei.columbia.edu/2013/04/29/contest-aims-to-bring-dark-data-into-digital-archives/>.

Morrissey, Shelia, et al. "Portico: A Case Study in the Use of the Journal Archiving and Interchange Tag Set for the Long Term Preservation of Scholarly Journals." *National Center for Biotechnology Information*. 2010. <http://www.ncbi.nlm.nih.gov/books/NBK47087/archive-content-access/access-to-archived-content> (accessed 2013).

Portico. *Portico- A Digital Preservation and Electronic Archiving Service*. 2005-2013. <http://www.portico.org/digital-preservation/>.

Portico. *Portico- A Digital Preservation and Electronic Archiving Service*. 2005-2013. <http://www.portico.org/digital-preservation/> (accessed September 20, 2013).

Royal Society Publishing. *Digital Preservation - Our Commitment to Archiving*. 2013. <http://royalsocietypublishing.org/site/librarians/DigitalPreservation.xhtml> (accessed September 2013).