

PROJECT REPORT

Digitally Endangered Species: The BitList

Sarah Middleton and William Kilbride

Digital Preservation Coalition, UK

Corresponding author: Sarah Middleton (sarah.middleton@dpconline.org)

The Digital Preservation Coalition (DPC) is a not-for-profit professional membership organisation that enables its members to deliver resilient long-term access to digital content and services, helping them to derive enduring value from digital assets and raising awareness of the strategic, cultural and technological challenges they face. The Coalition achieve its aims through advocacy, community engagement, workforce development, capacity-building, good practice and good governance.

On 30 November 2017, as part of its ongoing advocacy work and the first International Digital Preservation Day, the DPC published the first 'BitList.' This list presented the results of a crowd-sourcing exercise that sought to discover which digital materials the digital preservation community thinks are most at risk, as well as which materials are relatively safe thanks to digital preservation.

Keywords: digital preservation; digital curation; digital archiving; digital sustainability; data loss

Digitally Endangered Species: The BitList

The DPC recently published a list of the world's digitally endangered species (<https://www.dpconline.org/our-work/bit-list>), outlining the digital materials that the global digital preservation community thinks are most at risk. Nicknamed the 'BitList', the endeavour aimed to raise awareness about the most urgent and changing risks faced in digital preservation. The mainstream media rarely discusses digital preservation; when it does, it presents the topic in the context of a 'digital dark age' without reference to the tireless, valuable work undertaken by digital preservation professionals around the world since the problems with obsolescence were first identified. In the unlikely event that media outlets mention solutions, they tend to offer technical solutions to a socio-technical problem, thus missing significant parts of the challenge that pertain to people, processes, strategy and policy. In an effort to address this omission, the DPC conceived the BitList, with the goals of allowing more subtle messages about the nature of the threats faced by digital materials, creating a framework for celebrating genuine achievements, and spelling out the consequences of failing to address problems that need urgent attention.

In compiling the list, the DPC has acted with experience and expert knowledge on behalf of a wide cross section of the global digital preservation community. The Coalition enables members to deliver resilient long-term access to digital content and services through community engagement, advocacy, workforce development, research and good practice, capacity-building, and partnership.

Process

The BitList was modeled on the IUCN 'RedList,' which provides a dynamic and well-understood classification of wildlife species, from 'extinct' to 'least concern' through a spectrum of classes of endangerment. But whereas species of flora and fauna are well classified using a taxonomy that has existed for centuries, the digital domain is harder to pin down because it is a relatively new and constantly changing field of study. More accurately, it is hard to establish a taxonomy that does not pre-judge the nature of the threats that digital objects face. For example, using a registry like Pronom to map the digital universe would suggest file formats were the beginning and end of preservation risks. The same is true of media classifications. So instead of pre-judging the outcomes, the DPC issued a Call for Nominations to the digital preservation







Home > Our Work > Digitally Endangered Species

The 'Bit List' of Digitally Endangered Species

The DPC's 'Bit List' of Digitally Endangered Species is a crowd-sourcing exercise to discover which digital materials our community thinks are most at risk, as well as those which are relatively safe thanks to digital preservation. By compiling and maintaining this list over the coming years, the DPC aims to celebrate great digital preservation endeavors as entries become less of a 'concern,' whilst still highlighting the need for efforts to safeguard those still considered 'critically endangered.'

[Listen to William Kilbride talking to the BBC Click programme about the Bit List](#)

Click on each of the Risk Classifications to see the entries for each category.

<p>LOWER RISK</p>  <p>Digital materials are listed as <i>Lower Risk</i> when it does not meet the requirements for other categories but where there is a distinct preservation requirement. Failure or removal of the preservation function would result in re-classification to one of the threatened categories.</p>	<p>VULNERABLE</p>  <p>Digital materials are listed as <i>Vulnerable</i> when the technical challenges to preservation are modest but responsibility for care is poorly understood, or where the responsible agencies are not meeting preservation needs. This classification includes <i>Lower Risk</i> materials in the presence of aggravating conditions.</p>	<p>ENDANGERED</p>  <p>Digital materials are listed <i>Endangered</i> when they face material technical challenges to preservation or responsibility for care is poorly understood, or where the responsible agencies are poorly equipped to meet preservation needs. This classification includes <i>Vulnerable</i> materials in the presence of aggravating conditions.</p>
<p>CRITICALLY ENDANGERED</p>  <p>Digital materials are listed <i>Critically Endangered</i> when they face material technical challenges to preservation, there are no agencies responsible for them or those agencies are unwilling or unable to meet preservation needs. This classification includes <i>Endangered</i> materials in the presence of aggravating conditions.</p>	<p>PRACTICALLY EXTINCT</p>  <p>Digital materials are listed as <i>Practically Extinct</i> when the few known examples are inaccessible by most practical means and methods. This classification includes <i>Critically Endangered</i> materials in the presence of aggravating conditions.</p>	<p>CONCERN</p>  <p>Digital materials are listed as of <i>Concern</i> when an active member of the digital preservation community has expressed a legitimate concern but the concern has not yet been assessed by the BitList jury. They will be assessed for inclusion in the subsequent year.</p>

community which aimed to solve two problems: to make proposals about what should go on the BitList and to make sense of those proposals once received.


This call was met by universities, hospitals, national and local libraries, archives, broadcasters and technologists around the world, who submitted around 100 nominations in September 2017 following a planning, communications and lead-up process which began earlier in the year. An international jury of DPC members undertook a first review of the nominations in October 2017, ranking and commenting on the entries. The nomination process saw some duplication, overlap, and nesting, so the second phase of assessment consolidated and simplified the 100 or so entries to 20 distinct groups. Even then, there were more nominations than the DPC and the BitList jury had time to process.

Findings

Some familiar trends emerged in the nominations to the BitList (now released), such as portable magnetic media or proprietary software, which are well-documented concerns for the digital preservation community. The digital preservation community has also begun worrying more seriously about human deficiencies, business failure, and deliberate obfuscation. The efforts once spent addressing the risks of obsolescence or media degradation now also need to address poorly constructed rights management, political interests, failing markets, and simple human frailties. Digital preservation needs an equal investment in policy and regulation, which in turn means people – governments, organizations, companies, lawmakers – need to take greater responsibility for preservation and not simply frame it as a technical challenge.

The BitList is not presented as a hierarchical 'top ten.' Instead, broad groups of content have been associated with broad categories of risk:

- **PRACTICALLY EXTINCT:** Digital materials are listed as Practically Extinct when the few known examples are inaccessible by most practical means and methods.
- **CRITICALLY ENDANGERED:** Digital materials that face material technical challenges to preservation. There are no agencies responsible for them or those agencies are unwilling or unable to meet preservation needs.
- **ENDANGERED:** Digital materials that face material technical challenges to preservation, materials for which responsibility for care is poorly understood, or materials where the responsible agencies are poorly equipped to meet preservation needs.
- **VULNERABLE:** Digital materials for which the technical challenges to preservation are modest, but responsibility for care is poorly understood or the responsible agencies are not meeting preservation needs.

Pre-WWW ViewData and TeleText Services where no archival agency has captured and retained the signal		
Pre WWW television information services broadcast within the TV signal that allowed a degree of search and retrieval of up-to-date information, based on TeleText or ViewData technologies and variants.		
Action: Action Most Urgent (Action and reporting within 12 months when discovered) Assessment Completed: November 2017	Examples: AerTel, Electra, MetroText, Antiope-based systems	Aggravating Conditions: Lack of understanding structure of information silos; Lack or loss of documentation; Uncertainty about intellectual property rights Lack of funding or impetus
		
Critically Endangered: where good practice can be demonstrated.		
<small>Note: this category was originally proposed as 'Pre-WWW ViewData and TeleText Services' alone without the qualifier 'where no archival agency has captured and retained the content'. This qualification was introduced and the examples of Ceefax and Teletext Oracle also removed upon notification that an archival agency has been identified holding comprehensive examples of these materials. It is actively developing a solution to provide robust digital preservation and access. If this work succeeds then the entire category can be downgraded from 'Practically Extinct' at the next review; if not then the qualifiers may be removed and the 'Practically Extinct' classification confirmed.</small>		

- **LOWER RISK:** Digital materials that do not meet the requirements for other categories but have a distinct preservation requirement. Failure or removal of the preservation function would result in reclassification to one of the threatened categories.

Timelines for action are suggested for each category, and where good practice is evident, the risk reduces; in the presence of aggravating conditions, the risk increases. The category of greatest concern, 'Practically Extinct,' has been reserved for materials that present significant technical challenges and have no obvious archival home. There were only two items in that group: Pre WWW View Data services (also known as Teletext) and Pre WWW Videotex Data Services (such as Bulletin Boards). This number might seem to underestimate the overall potential for digital loss. However, in the next category, Critically Endangered, there are ten items that, in the presence of aggravating conditions (e.g. a lack of understanding, structure of information silos, lack or loss of documentation, uncertainty about intellectual property rights, and/or lack of funding or impetus), also face near extinction. Recognizing that 'Politically Sensitive Material' or 'Unpublished Research Outputs' (i.e. research data) are in this category acknowledges that a large proportion of the digital estate is at high risk.

Since the publication of the BitList, the DPC has already received follow-up information regarding items on the list such as TeleText services, which was categorized as 'Practically Extinct.' The term 'Practically Extinct' had been employed deliberately to invite correction from the digital preservation community and to draw attention to those places where content may yet be saved. Although extensive research had been undertaken to assess the availability of this content, after publication of the list the DPC was alerted to the efforts of an underfunded research project that had been trying to migrate the TeleText component out of the broadcast video recordings and into a searchable digital interface. The qualification 'where no archival agency has captured and retained the content' has now been introduced, and the examples of Ceefax and Teletext Oracle have been removed. If this work succeeds, the entire category can be downgraded from 'Practically Extinct' at the next review; if not, the qualifiers may be removed and the 'Practically Extinct' classification confirmed.

Conclusions and next steps

By compiling and maintaining the BitList over the coming years, the DPC will be able to celebrate digital preservation successes whilst still highlighting the need for efforts to safeguard those still considered 'Critically Endangered.' The current BitList will form the basis of the next edition, which is anticipated for review and re-release in November 2018 to coincide with this year's International Digital Preservation Day. The DPC invites comments and criticisms of the BitList. And if there are items that have been overlooked, then there is always space for a second edition.

Competing Interests

The authors have no competing interests to declare.

How to cite this article: Middleton, Sarah and William Kilbride. 2018. Digitally Endangered Species: The BitList. *KULA: knowledge creation, dissemination, and preservation studies* 2(1): 12. DOI: <https://doi.org/10.5334/kula.20>

Submitted: 05 February 2018

Accepted: 18 May 2018

Published: 29 November 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



KULA: knowledge creation, dissemination, and preservation studies is a peer-reviewed open access journal published by Ubiquity Press

OPEN ACCESS The Open Access logo, which is a stylized padlock icon with an open keyhole.