



# The BitList 2021

The Global List of Digitally  
Endangered Species

Third Edition

Revised

November 2021

## 1. Executive Summary

The Global List of Digitally Endangered Species – *The BitList* – offers an accessible snapshot of the concerns expressed by the global digital preservation community with respect to the risks faced by diverse types of digital content in varied conditions and contexts. It provides an elementary assessment of the imminence and significance of the dangers faced by different, and at times overlapping classifications of digital materials. By identifying the urgency of action and significance of content, *The BitList* draws attention to those digital materials that, in the view of the global digital preservation community, require urgent action to remain viable.

The list is derived directly from the practical experience of professionals with the responsibility to maintain access to content over time: it is their voice which *The BitList* represents. They come from around the world and from many different sectors. It is not a top-down or theoretical exercise, nor does it serve a political or commercial interest. At a fundamental level, items appear on the list because an established and experienced professional within the digital preservation community has struggled to preserve access to this content and has called for it to be included.

The categories and classifications of content are broad so that the list can be digested quickly. This accessibility comes at a cost to specificity. It is a reference set against which any digital object can be compared. The urgency of action or risks faced are amplified by the presence of aggravating factors; and they are ameliorated in the presence of good practice. Entries overlap. Any given digital object may appear under multiple headings depending on technology, resourcing or organizational context. These overlapping classifications mean that objects may be at greater risk than initially suggested and that actions to tackle the risks are potentially more complex. A condensed action plan is suggested for every entry.

First and foremost, *The BitList* is an advocacy tool. It emerged as a recommendation from the DPC's Advocacy and Community Engagement Sub-Committee and exists to inform priorities in the allocation of resources and to support policy development where needed. It is intended as an ongoing framework of assessment that highlights risks and provides a basis for the celebration insofar as challenges, once identified, will be resolved and reported in subsequent years. The framework is also intended to become more specific over time and thus more direct in recommendations.

*The BitList* has three main audiences and three related functions:

- As an advocacy tool, *The BitList* seeks to influence the technology sector and senior leaders within corporations and agencies of all kinds, giving them an honest but accessible account of the threats the digital preservation community perceives and the progress it has made. It seeks to influence them to invest credibly and plan with a more informed sense of the risks that are faced by digital materials in the longer term;
- As a practical comment on the challenges faced across the digital preservation community, *The BitList* provides a rudimentary but practically informed development roadmap, whether for researchers in academic institutions or commercial and semi-commercial agencies who seek to bring products to the market. It invites them to consider and, where possible, resolve the challenges that are identified here and offer credit as solutions are progressed;
- As a state-of-the-art report, *The BitList* provides introductory and current guidance for the digital preservation community. This is especially useful for new entrants as well as teachers, so that they are prepared for the challenges and opportunities that arise in the practice of digital preservation. It also supports professionals of long standing that may be approaching new challenges or content for the first time.

*The BitList* was first published in 2017, and since then, has undergone a comprehensive review every two years with an interim progress report and commentary in alternate years. This schedule explicitly complements the biannual cycle of the Digital Preservation Awards.

A major review of both newly nominated and existing BitList entries was conducted in 2021 in line with this schedule. As done in previous iterations of *The BitList*, the comprehensive review brought together the different backgrounds, knowledge and expertise of a panel of experts and practitioners within the digital preservation community. These members of *The BitList* 2021 Jury were critical for ensuring the robustness of the list. They gave their feedback as well as recommended names of potential experts and examples of case studies to add to the robustness of the revised list of digitally endangered species. In one form or another, *The BitList* 2021 has been passed through the following stages of review between open nomination and publication:

- Eligibility and credibility check
- Individual assessment of imminence, significance, impact, and effort required (Scoring Round 1)
- Expert and Jury panel review (Jury Meeting 1)
- Individual assessment or re-assessment of entries with diverging scores in preparation for Jury discussion (Scoring Round 2)
- Expert and Jury panel review and vote (Jury Meeting 2)
- Peer review by Jury, DPC Staff and recommended experts

*The BitList* 2021 offers an extensive but not exhaustive list of digital materials at risk. Each entry on the list was addressed during the review process, reconsidered by panel experts in light of new evidence or changed circumstances, and then revised, rescored or restructured where needed to amplify the current risks faced by the digital content and possible steps for reducing the likelihood or impact of loss within a suggested timeframe.

The DPC, which manages and publishes *The BitList*, maintains ‘neutrality in respect to solutions, approaches, sectors and vendors.’ This position is embedded by constitution, value and practice and maintained scrupulously throughout the DPC’s operations. Thus, the recommendations and classifications have been assembled independently of the interests of vendors or solution providers.

While this year’s edition of *The BitList* is considerably more robust in content and process than before, it is interesting to note that the total number of entries is actually lower than in the previous edition—there are now 73 entries on the list compared with 74 in the previous year!

- *Lower Risk*: 0
- *Vulnerable*: 8
- *Endangered*: 27
- *Critically Endangered*: 32
- *Practically Extinct*: 6

This is not to say there are fewer items at risk, but rather that the deeper assessments of the entries by the Jury led to several entries being merged or better situated as case examples within a broader grouping. The hope is that the risks associated with their digital content are easier to locate and communicate from this deeper assessment of the entries.

By its nature, *The BitList* is always a provisional statement. It is published and reviewed with the understanding that new risks are continuously arising every day and (inevitably) between editions. The extent of the digital domain, the complexity of the threats, and the sophistication of emerging solutions mean that no process could ever fully capture the risks and challenges faced by digital content around the world. The Jury also recognizes that differences in emphasis and subtleties of local context may well have

been overlooked, and that material changes may have occurred during the process of compilation, which should be taken into consideration for the next revision. We welcome corrections and suggestions on how the list could be improved, encourage you to act and share your own case studies or examples, invite you to participate in the global community which is developing good practice around digital preservation in the enhancement of the list for its next interim review scheduled for November 2022.

## Table of Contents

1. Executive Summary .....	1
2. About the Digital Preservation Coalition.....	7
3. Acknowledgements: The BitList Jury .....	8
4. Director's Introduction to The BitList 2021 .....	9
5. Interpreting and Using The BitList .....	13
6. Explanation of Classifications .....	15
Lower Risk.....	15
Vulnerable .....	15
Endangered .....	15
Critically Endangered.....	15
Practically Extinct .....	15
7. Lower Risk.....	16
8. Vulnerable .....	17
Cloud Storage .....	18
Current Hard Disk Technologies.....	20
Local Network Storage .....	21
PDF/A.....	22
Pension, Mortgage and Insurance Records.....	23
Published Research Papers.....	25
Recently Commissioned or Completed Media Art .....	26
Research Data Published through Repositories .....	28
9. Endangered.....	29
3D Digital Engineering Drawings .....	30
Born Digital Photographs and Video shared via Social Media or Uploaded to Cloud Services.....	32
Cloud-based Services and Communications Platforms .....	34
Completed Investigations based on Open Source Intelligence Sources .....	36
Content on Cloud Video Services Produced by the Service Provider.....	37
Contractual Documents and Related Records.....	39
Corporate Records of Long Duration on Network Drives, Intranets and EDRMS .....	41
Current Portable Magnetic Media.....	43
Current Portable Optical Media .....	44
Current Portable Solid-State Media .....	45
Digital Music and Ephemera Shared on Social Media.....	46

Digital Radio Recordings.....	48
Digital Recordings Published via Cloud-based Music Sharing Platforms .....	50
Digitally Published Sheet Music.....	52
Electronic Hospital and Medical Records .....	53
Email .....	55
Legacy Media Art.....	57
Legacy Video Files.....	59
Master Digital Music and Sound Recordings.....	61
Orphaned Works .....	63
PDF other than PDF/A .....	65
Pre-Production TV and Movie Materials .....	67
Proceedings in Court .....	68
Published Research Data Appended to Journal Articles .....	69
Recordings of Video Game Play Uploaded to Online Platforms .....	71
Semi-Published Research Data.....	73
Virtual Reality Materials and Experiences.....	75
10. Critically Endangered.....	77
Community-generated Content in Arts and Heritage .....	78
Consumer Social Media Free at the Point of Use.....	80
Correspondence and Records of Research.....	82
Data Posted to Defunct or Little-used Social Media Platforms.....	84
Digital Archives from Public Enquiries and Commissions .....	86
Digital Archives of Community Groups .....	88
Digital Archives of Music Production .....	90
Digital Evidence and Records of Investigation Prior to Court .....	91
Evidence in Court.....	92
Family or Personal Records .....	94
Grey Literature .....	96
Legacy Research Web Collections .....	98
Massively Multiplayer Online Gaming Platforms and Experiences.....	99
Media Art by Deceased Artists or Defunct Workshops.....	101
Media Inside Paper Files.....	103
Non-current Hard Disk Technologies .....	104
Non-current Portable Magnetic Media .....	105
Non-current Portable Optical Media.....	106
Non-current Portable Solid-State Media.....	107

Old or Non-current Offline Video Games .....	108
Old or Non-current Video Games.....	110
Open Source Intelligence Sources of Current Conflicts.....	112
Pre-WWW ViewData and Teletext Services where no Archival Agency has Captured and Retained the Signal .....	113
Politically Sensitive Data.....	115
Records of Local Government .....	117
Records of Non-Governmental Agencies .....	119
Records of Quasi Non-Governmental Agencies .....	121
Smart Phone Apps .....	123
Smart Phone Gaming.....	125
Supporting Digital Materials for Museums and Galleries .....	127
Unpublished Research Data from Government Researchers.....	129
Web Domains with No Legal Deposit .....	130
11. Practically Extinct.....	131
Adobe Flash Animations and Interactive Applets .....	132
Legacy Interfaces and Services Offered Online by Major Companies.....	133
Non-standard Public Records .....	135
Older Open Source Intelligence Sources .....	137
Pre-WWW Videotex Data Services and Bulletin Board Services .....	139
Unpublished Research Data .....	140
12. Appendix.....	141

## 2. About the Digital Preservation Coalition

The Digital Preservation Coalition (DPC) is an international charitable foundation which supports digital preservation, helping its members around the world to deliver resilient long-term access to digital content and services through community engagement, targeted advocacy work, training and workforce development, capacity building, good practice and standards, and through good management and governance. The DPC exists to secure our digital legacy.

The DPC's work, including *The BitList* is shaped by our values. In all that it does, the DPC:

- Maintains neutrality in respect to solutions, approaches, sectors and vendors
- Amplifies the needs and successes of our members
- Is open to all stakeholders
- Maintains sound stewardship of our shared resources
- Responds to the needs of members in the delivery of services
- Is authoritative, current, and concise in all our publications and communications
- Is respectful, welcoming, inclusive, and transparent in all our dealings

Members and colleagues who engage with the DPC recognize these values in their experience of the DPC and may be asked to adopt them when working with us or interacting through us.

The scope of the DPC and its activities are defined in five ways:

- **By our members:** we invest considerable energy in framing a program that responds to members' needs. As the membership changes so this scope will change through time.
- **By topic:** we define digital preservation as the managed activities necessary to ensure continued access to digital materials for as long as necessary, including all the actions required to maintain access beyond the limits of media failure, technological obsolescence or community change. We engage in and with any and all of the people, tools, services, agencies and activities that aid this purpose.
- **By sectors:** we are a cross-sector, inter-disciplinary body, open to all who need to ensure continuing access to digital content, irrespective of purpose.
- **By locale:** digital preservation is a global challenge so the DPC makes a global offer. We welcome memberships, partnerships and collaborations with agencies and individuals around the world. Our origins are as a joint endeavour between agencies in the UK and Ireland where the bulk of our members are still situated. At the start of this planning period, DPC is active in 10 countries and 3 continents. In the delivery of this strategic plan, we will extend our global offer and prepare our transition to a global foundation.
- **By our values:** we maintain our neutrality with respect to solutions, approaches and vendors because this protects the independence and value of our activities. But neutrality does not imply exclusion from meaningful engagement.

For more about the Digital Preservation Coalition, including how to join, see: <https://www.dpconline.org/>



### 3. Acknowledgements: The BitList Jury

The DPC gratefully acknowledges the support of the worldwide digital preservation community in the compilation of *The BitList*.

In particular, we are grateful to the expert panel who supported its creation and assessed the entries received in 2021:

- Paul Stokes (Jisc), *Chair of the Bit List Jury*
- Ruth Cammies (Open University),
- Elizabeth England (National Archives and Records Administration),
- Tom Ensom (Tate),
- Emma Hancox (University of Bristol),
- Patricia Herterich (Digital Curation Centre),
- William Kilbride (DPC),
- Micky Lindlar (Technische Informationsbibliothek (TIB)),
- Kirsty Lingstadt (formerly University of Edinburgh, now University of York),
- Annette Mills (National and State Libraries Australia),
- Alexandra Mitchell (University of Salford),
- Svenia Pohlkamp (nestor),
- Arran Rees (University of Leeds),
- Sebastian Roncin (Bacardi-Martini),
- Anthea Seles (International Council on Archives),
- Melanie Swalwell (Swinburne University of Technology),
- Sara Day Thomson (University of Edinburgh),
- Tamara van Zwol (Digital Heritage Network).

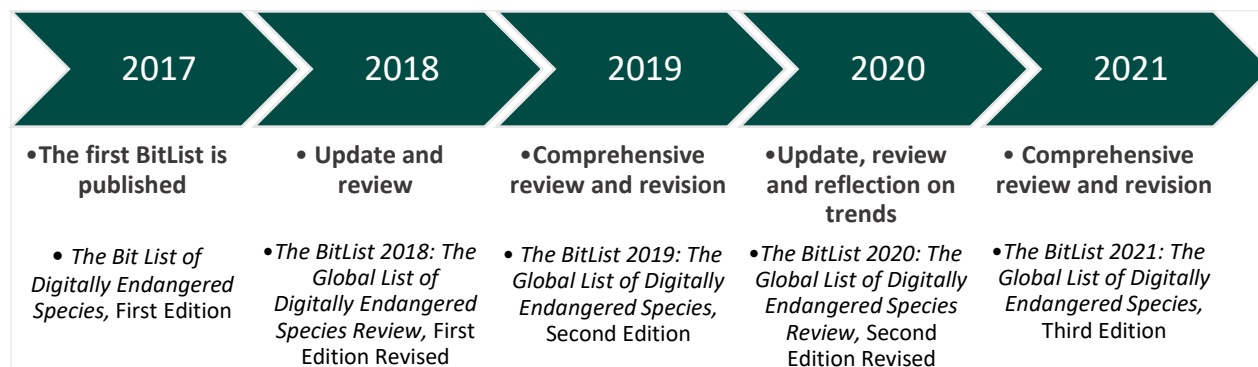
The 2021 Revision of the Bit List was completed by the DPC, led by Amy Currie and Sarah Middleton with reviews by and feedback from the Bit List Jury, the DPC Advocacy & Community Engagement Subcommittee, and DPC staff.

We are grateful for the additional expertise provided generously by subject matter experts.

## 4. Director's Introduction to The BitList 2021

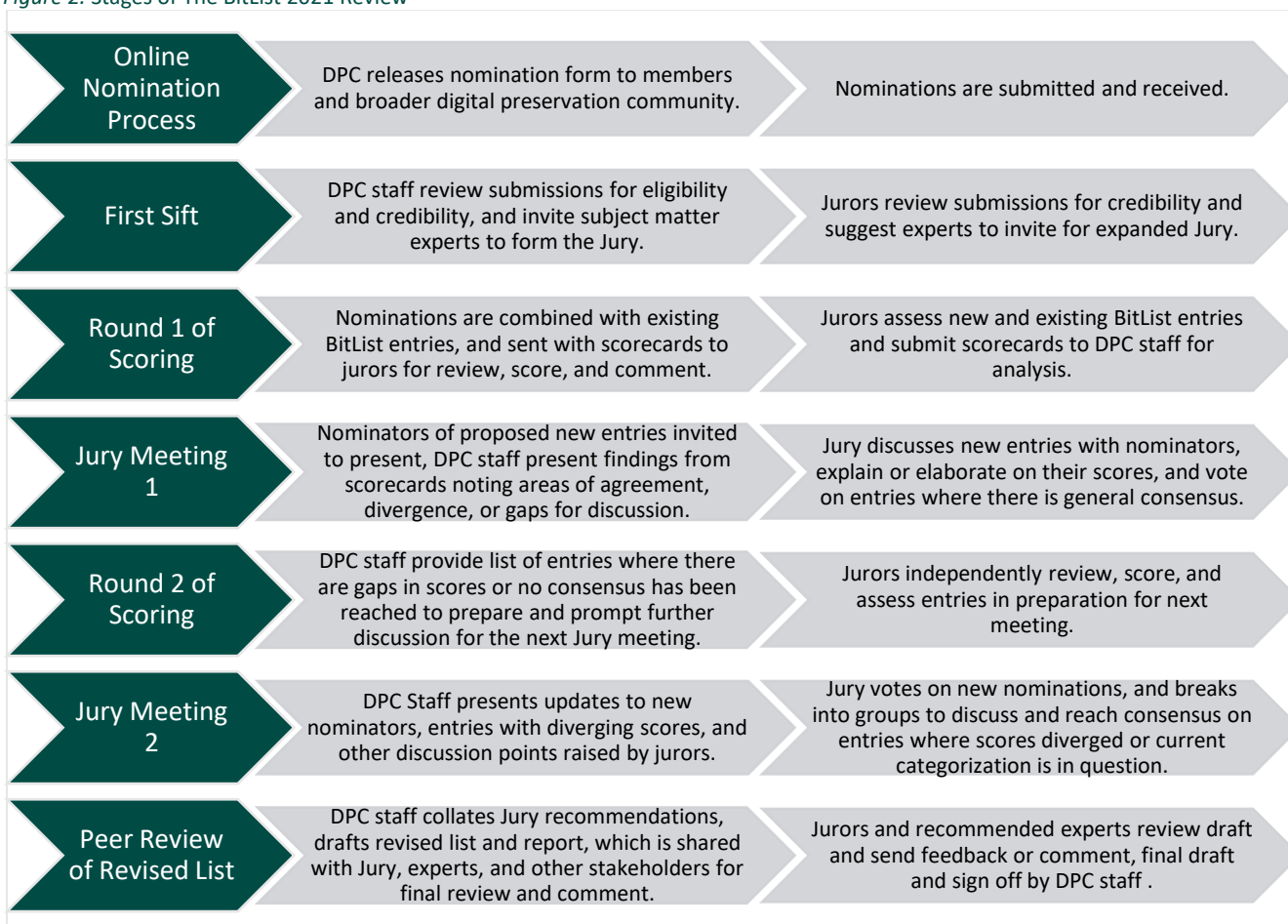
*The BitList 2021* is the second comprehensive review of the list since its initial publication in 2017. It builds on the work of previous *BitList* juries, in particular the comprehensive review and revisions for *The BitList 2019* and *The BitList 2020's* introduction of identified risk trends.

Figure 1. History of the BitList



Each entry on *The BitList 2021* was assessed by jury panel experts as part of the review process and subsequently revised, rescored or restructured where needed to better reflect the risks associated with the digital material and developments or updates since the last review. The review process took place from June to November 2021, and in one form or another, passed through the following stages:

Figure 2. Stages of The BitList 2021 Review



The resulting *BitList 2021* offers an extensive but not exhaustive list of digital materials at risk.

Readers may notice that the length of the report is considerably longer than ever before. For this reason, we offer a few highlights to introduce here at the start, especially when reflecting on how the wider context has changed over the last two years since the last comprehensive review.

### Major Revisions to the List of Digitally Endangered Species

To begin, while it is considerably more robust in content and process than prior years, it is interesting to point out that the total number of entries on the list is marginally lower than in the previous edition (73 compared to 74 in the second edition).

This is not to say there are fewer items at risk than before, but rather that the deeper assessments of the entries during the review led to several entries being merged or changed from standalone entries to case examples within a broader grouping. These include:

- *A new merged entry.* ‘Cloud-based Services and Communications Platforms’ ([p. 34](#)), formed from two previous entries of ‘Consumer cloud-based utilities’ and ‘Premium or institutional social media services.’
- *A case example formed from a previous entry.* The previous entry, ‘Born Digital Images Held Offline on Portable Storage Devices’, applies to multiple broader portable storage media entries and could easily be overlooked in that respect. Therefore, it was moved from a standalone entry in the list and incorporated into the BitList overview page on Portable Media to highlight the risk of born digital images in and across different types of portable storage media.
- *Two case examples formed from two previous entries.* Rather than sit as standalone entries, the Jury recommended two previous entries, ‘Maritime Archaeological Archives’ and ‘Geomagnetic Data and Software,’ become case examples under the broader grouping of research outputs to better illustrate how research data is complex and has specific requirements for documentation that may only be known to subject matter experts. They were therefore incorporated into the BitList overview page on Research Outputs.

Conversely, the Jury review resulted in the disaggregation of one entry into two subset entries. The 2019 ‘Proceedings and Evidence in Court,’ a single entry in 2020, is now represented by two entries: ‘Proceedings in Court’ ([p. 68](#)) and ‘Evidence in Court,’ ([p. 92](#)) which have different risk profiles requiring similar but distinct preservation actions.

There were entries where the scope was broadened or narrowed. The previous ‘Unpublished Research Data from US Govt Researchers’ broadened to include those in various national contexts (now titled ‘Unpublished Research Data from Government Researchers,’ [p. 129](#)). The previous ‘Digital Materials in Museums and Galleries’ entry narrowed in scope to emphasize supporting materials within Museums and Galleries with higher risk profiles (now titled ‘Supporting Digital Materials for Museums and Galleries,’ [p. 127](#)). The previous ‘Offline Gaming’ entry also narrowed to address aggravating conditions and higher risk profiles for offline games which are old or non-current (now titled ‘Old or Non-current Offline Games and Gaming,’ [p. 108](#)).

### The Addition of Two New Entries

As many readers may be most excited to hear, two new entries were added through the open online nomination process: ‘Virtual Reality Materials and Experiences’ ([p. 75](#)) and ‘Smart Phone Gaming’ ([p. 125](#)). As detailed in their entry descriptions, these new entries bring to attention their historical and cultural significance but, more importantly, the complexity and interconnectedness of their various technologies that make their preservation challenging, and the impact of loss to access to them for wide-ranging uses across many sectors.

### Increasing Risk Trends and Classifications

Finally, it is important to recognize where the context for digital preservation – as for much else – has been dramatically disrupted by the Coronavirus Pandemic, social and political unrest, and changing global climate calls for action. *The BitList 2020* report observed how, despite the heroic efforts of the digital preservation community, there was a noticeable trend towards increased risk in 2020. That trend has continued into 2021. Circumstances relating to the Covid pandemic and continued political and economic volatilities over the last year have impacted many entries. The Jury identified 36 items on *The BitList 2021* where there is a trend towards greater risk (compared with 17 in 2020) with 7 of those 36 items assigned a higher risk classification.

The seven entries with a trend towards greater risk with changed classification were:

- ‘Adobe Flash Animations and Interactive Applets’ changed from *Critically Endangered* to *Practically Extinct* ([p. 133](#))
- ‘Consumer Social Media Free at the Point of Use’ changed from *Endangered* to *Critically Endangered* ([p. 80](#))
- ‘Records of Quasi Non-Governmental Agencies’ changed from *Endangered* to *Critically Endangered* ([p. 121](#))
- ‘Massively Multiplayer Online Gaming Platforms and Experiences’ changed from *Endangered* to *Critically Endangered* ([p. 99](#))
- ‘Supporting Digital Materials for Museums’ (formerly ‘Digital Materials in Museums and Galleries’) changed from *Endangered* to *Critically Endangered* ([p. 127](#))
- ‘Content on Cloud Video Services Produced by the Service Provider’ changed from *Vulnerable* to *Endangered* ([p. 37](#))
- ‘Digital Recordings Published via Cloud-based Music Sharing Platforms’ changed from *Vulnerable* to *Endangered* ([p. 48](#))

The other 29 entries with a trend towards greater risk (but no change in classification) were:

- 3D Digital Engineering Drawings ([p. 30](#))
- Born Digital Photos and Video Shared on Social Media or Uploaded to Cloud Service ([p. 32](#))
- Cloud-based Services and Communications Platforms (new merged entry combining previous ‘Consumer cloud-based utilities’ and ‘Premium or institutional social media services’ entries) ([p. 34](#))
- Cloud Storage ([p. 18](#))
- Completed Investigations based on Open Source Intelligence Sources ([p. 36](#))
- Corporate Records of Long Duration on Network Drives, Intranets and EDRMS ([p. 41](#))
- Current Hard Disk Technologies ([p. 20](#))
- Data Posted to Defunct or Little-used Social Media Platforms ([p. 84](#))
- Digital Archives of Community Groups ([p. 88](#))
- Digital Evidence and Records of Investigation Prior to Court ([p. 91](#))
- Digital Music and Ephemera Shared on Social Media ([p. 46](#))
- Electronic Hospital and Medical Records ([p. 53](#))
- Evidence in Court (new subset entry from previous ‘Proceedings and Evidence in Court’) ([p. 92](#))
- Legacy Media Art ([p. 57](#))
- Legacy Research Web Collections ([p. 98](#))
- Legacy Video Files (formerly ‘Video Files’) ([p. 59](#))
- Media Art by Deceased Artists or Defunct Workshops ([p. 101](#))
- Media Inside Paper Files ([p. 103](#))

- Non-current Hard Disk Technologies ([p. 104](#))
- Old or Non-current Offline Video Games (formerly ‘Offline Gaming’) ([p. 108](#))
- Orphaned Works ([p. 63](#))
- Politically Sensitive Data ([p. 115](#))
- Pre-production TV and Movie Materials ([p. 67](#))
- Recently Commissioned or Completed Media Art ([p. 26](#))
- Recordings of Video Game Play Uploaded to Online Platforms ([p. 71](#))
- Records of Local Government ([p. 117](#))
- Records of Non-Governmental Agencies ([p. 119](#))
- Unpublished Research Data from Government Researchers (formerly ‘Unpublished research data from US Govt researchers’) ([p. 129](#))
- Web Domains with No Legal Deposit ([p. 130](#))

### Decreasing Risk Trends and Classifications

The review did not lead to all bad news. There is good news to share. Eleven entries had a trend towards reduced risk (compared to one in 2020), with one changing to a *Lower Risk* classification.

The entry for ‘Pre-WWW ViewData and Teletext Services where No Archival Agency has Captured and Retained the Signal’ was changed from *Practically Extinct* to *Critically Endangered* ([p. 113](#)). This change took into account a number of important developments reported since this it was first added to the *BitList* in 2017, raising hope that collections can be recovered and re-used under certain circumstances. A trend towards decreased risk was noted in 2020, based on live-capture of broadcast output at the BBC and British Film Institute, and teletext graphics characters were added to the Universal Coded Character Set, making it easier to transmit and archive teletext and legacy computer graphics for archiving and preservation. The 2021 Jury agreed on a trend towards reduced risk. While it is not yet fully clear how such a signal could be made searchable or made available at scale, research has progressed with different threads of research and enthusiast and community led initiatives being brought together in the development of preservation tools and techniques. In light of these developments and active research and recovery efforts, the Jury supported a change from the *Practically Extinct* to *Critically Endangered* classification.

The other ten entries with a trend towards reduced risk (but no change in classification) were:

- Current Portable Magnetic Media ([p. 43](#))
- Email ([p. 55](#))
- Grey literature ([p. 96](#))
- PDF/A ([p. 22](#))
- PDF Other than PDF/A ([p. 65](#))
- Published Research Papers ([p. 25](#))
- Published Research Data Appended to Journal Articles ([p. 69](#))
- Research Data Published through Repositories ([p. 28](#))
- Semi-Published Research Data ([p. 73](#))
- Unpublished Research Data ([p. 141](#))

The hope is that these additions and revisions presented in *The BitList 2021*—both good and bad—better reflect the significance of content and impact of loss for the materials at the present time and prompt a further call to action for them to remain viable.

*William Kilbride, November 2021*

*Executive Director, The Digital Preservation Coalition*

## 5. Interpreting and Using The BitList

The BitList is first and foremost an advocacy tool. It describes a range of digital materials in varied organizational settings which, in the experience of the global digital preservation community, face distinct and imminent challenges. These challenges may be as much to do with accountability, policy or business process as technological obsolescence or media decay. By identifying them and by providing elementary recommendations about how the risks can be tackled, the DPC seeks to provide generic, impartial and international support to specific preservation actions and policies in any context.

Although all digital materials fall within the scope of the BitList, it is not a complete account of digital materials at risk: only those items which members of the community recognize as being at risk are included. Consequently, the fact that a data set is not listed should not be taken as evidence that it is not at risk: simply that the community which has compiled the list has not encountered any explicit risks or has no experience with these materials. Equally, the fact that an item has been identified as at risk is some small proof of effort, however weak, to secure long-term viability. In most cases, entries on the list are broadly defined, summarizing significant variability in specific cases. Many items on the list overlap, amplifying or lessening the urgency for action as appropriate.

Each item on the list is given a short title and a longer description. It is described in general terms, then a series of examples are given. The examples are illustrative not exhaustive, and in many cases, they are also broadly defined which in turn have many instances and examples. The examples typically include specific submissions made in the open nomination process and examples arising from Jury discussions.

Users of *The BitList* are encouraged to assess whether any digital object in their possession, or which they intend to create, or for which they have a current or imminent preservation responsibility, is a specific example of the item described and whether it aligns with one of the examples given.

Each item includes examples of *Aggravating Conditions* which amplify the risks a digital object faces, and *Good Practice* that would reduce the risk. These are also implied recommendations for addressing and reducing risks to be followed in the timescale indicated. In most cases, a fuller assessment is also suggested. By implication, the actions that would arise from such an assessment are not likely to be trivial. The Jury has attempted to provide a simple assessment of how much work it would be to improve the situation and their perception of how wide the impact of loss would be. Finally, detailed comments from the Jury have been included where available.

The Jury paid particular attention to the risk classification and trend. Items were given a provisional ranking by jurors during the first round of scoring, with entries requiring additional expert advice identified so that recognized subject matter experts could be invited to offer feedback on an item in more detail. The review process also saw the elimination of some entries, as well as the merger of duplicate entries and disaggregation of compound entries into smaller groups, following Jury discussion. Every entry here is on the basis of a consensus decision, but in a small number of cases, the decision was unanimous. For the sake of transparency, The BitList reports those occasions where unanimity was achieved as this materially affects how recommendations are deployed.

*The BitList 2021* has been compared to earlier editions of the list, and general comments have been added about trends and activities that have impacted each item now and in previous years.

Recognizing that entries are very broadly defined, digital materials can be at more or less risk depending on local circumstances. There is a greater risk, and therefore greater urgency to act, in the presence of aggravating conditions which can be delineated. So, while an entry may be classified as *Vulnerable* in generic terms, any example of that entry may reasonably be described as *Endangered* or *Critically Endangered* in the

presence of aggravating conditions. Conversely, in the presence of good practice, specific digital materials may move from *Endangered* to *Vulnerable* or *Lower Risk*.

*The BitList* offers a provisional statement with the recognition that the extent of the digital domain, the complexity of the threats, and the sophistication of emerging solutions mean that no process could ever fully capture the risks and challenges faced by digital content around the world. It is published and reviewed with the understanding that new risks are continuously arising; every day and (inevitably) between editions of *The BitList*. The Jury recognizes that differences in emphasis and subtleties of local context may well have been overlooked, and that material changes may have occurred during the process of compilation.

*The BitList* is designed to be collaborative, iterative and provisional. Thus, if readers are aware of significant digital collections that do not match up with any of the broad examples given but are at material risk, they are encouraged to draw these to the attention of the Jury through the DPC's Head of Advocacy and Community Engagement. These will be reviewed in time for publication of the next scheduled update in November 2022. Where digital materials face an imminent extinction event before that, their evaluation may be accelerated and an addendum published to *The BitList* in order to provide the timely, impartial and expert advocacy that may be required.

The Jury and DPC also recognize the strengths and limitations of the methods employed for the review. The scoring of significance, impact, the inevitability of loss, and imminence of action using numerical scales helped identify entries with general consensus, guide discussion and supplement arguments. There was the hope that the scoring metrics for the scales would be straightforward, and the gathering and analysis of numerical scores would help with standardization. This was indeed helpful when looking at the scores relating to the recommended risk classification; with the exception of two entries, all the scores submitted by jurors fell under one standard deviation.<sup>1</sup> However, the majority of submitted scores for significance and impact varied more widely. Jurors raised questions and comments on different interpretations of terms used in the scoring metrics and also on how well scores could be uniformly applied across content-, media-, or context-specific entries.<sup>2</sup> Jury meeting discussions expanded on the different interpretations to explain their reasoning and support arguments but, in light of the points raised, it is recommended that these comments be revisited and re-examined to improve the methods employed for the next comprehensive review scheduled for 2023.

---

<sup>1</sup> The two were 'Digital Archives of Music Production' (Standard Deviation of 1.22) and 'Maritime Archaeological Archives' (Standard Deviation of 1.00).

<sup>2</sup> Only five entries had a Standard Deviation below 1.00 for 'significance', and only five below 1.00 for 'impact'. See Appendix for a summary of the scoring criteria and comments raised by Jurors.

## 6. Explanation of Classifications

### Lower Risk

Digital materials are listed as *Lower Risk* when they do not meet the requirements for other risk classifications but where there is a distinct preservation requirement. Failure or removal of the preservation function would result in assignment to one of the more threatened classifications.

### Vulnerable

Digital materials are listed as *Vulnerable* 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.

### Endangered

Digital materials are listed as *Endangered* 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 *Vulnerable* materials in the presence of aggravating conditions.

### Critically Endangered

Digital materials are listed as *Critically Endangered* 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 *Endangered* materials in the presence of aggravating conditions.

### Practically Extinct

Digital materials are listed as *Practically Extinct* when examples cannot be identified or are inaccessible by most practical means and methods. It does not assume that the material is lost, but rather that loss is imminent and immediate action is required to avoid loss. It includes material where recovery is possible in very small samples but is impractical or has not been demonstrated at scale.



## 7. Lower Risk



Digital materials are listed as *Lower Risk* when they do 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.


There are no entries in The BitList 2021 which meet this description. However, digital materials described as *Vulnerable* become *Lower Risk* where good practice is applied. Suggestions of good practice have been made.

## 8. Vulnerable



Digital materials are listed as *Vulnerable* 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 would include *Lower Risk* materials in the presence of aggravating conditions; and endangered materials in the presence of good practice.


<b>Cloud Storage</b> Materials routinely copied or backed up to an independently managed, off-site data storage facility and are able to be restored under contractual terms.		
Group: Storage	<b>Trend in 2021:</b>	
Added to List: 2019	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Remote network storage provided by a third-party service under contracts, such as DropBox, Amazon, Microsoft Azure, Dell EMC, Google Cloud Platform, Google Drive, IBM, Intel, Rackspace, Iron Mountain, SAP, and others		
<b>'Endangered' in the Presence of Aggravating Conditions</b> Encryption; lack of routine maintenance; lack of storage replication; over-dependence on a single supplier; insufficient documentation; lack of local alternative; political or commercial instability; overly aggressive compression; poor information security; lack of transparent integrity-checking; lack of strategic investment; lack of migration plan; lack of exit strategy; unenforceable penalties; unstable pricing; unpredictable removal costs		
<b>'Lower Risk' in the Presence of Good Practice</b> Backup to different technology; backup to diverse locations; documentation of assets; integrity checking; preservation planning; export functionality; resilient to hacking; version control; resilient funding; technology watch; enforceable contract; disaster planning and documentation, stable pricing; budgeted removal costs.		
<b>2021 Review</b> This entry was added in 2019 to ensure that the range of media storage is properly assessed and presented. The 2021 Jury noted increased risk in light of greater reliance on the cloud and localized disruptions to cloud services over the pandemic. The 2021 trend towards greater risk is based on the wider (global) dependence on these services, especially Google Drive, for record-keeping and business workflows. The impact of loss increases with more reliance on cloud services leading to greater risk; however, this should not deter people from using cloud storage. Therefore, the <i>Vulnerable</i> classification stands so long there are safeguards in place.		
<b>Additional Comments</b> The history of digital preservation suggests that the risk of vendors going out of business or shutting down services is the key issue here, over and above any specific technical solutions or risks. <i>Case Studies or Examples:</i> <ul style="list-style-type: none"> <li>Case of a cloud storage provider who suffered major data loss (or its clients suffered data loss) due to a fire in its data centre. Those clients suffered most who did not include</li> </ul>		


geographically redundant storage in the contract with the storage provider as this was more expensive. See Rosemain, M. and Satter, R., (2021) Millions of websites offline after fire at French cloud services firm. *Reuters*. 10 Mar. Available at:


<https://www.reuters.com/article/us-france-ovh-fire-idUSKBN2B20NU>.

- Case of fired credit union employee accessing the financial institution's computer systems without authorization and destroying over 21 gigabytes of data via remote network storage <https://www.bleepingcomputer.com/news/security/fired-ny-credit-union-employee-nukes-21gb-of-data-in-revenge>


See also: Digital Preservation Coalition (2016) 'Storage' in Digital Preservation Handbook (Second Edition), online at <https://dpconline.org/handbook/organizational-activities/storage>.

<b>Current Hard Disk Technologies</b>		
Materials saved to storage devices with a variety of underlying magnetic or solid-state (flash) technologies that are hardwired into a computer still under warranty or supported: typically hard disks that are less than five years old.		
Group: Storage	<b>Trend in 2021:</b>	Consensus Decision
Added to List: 2019	<b>Trend towards greater risk</b>	Previously: Vulnerable
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Direct Attached Storage (DAS) such as magnetic or solid-state drives integrated into individual laptops or workstations and into smaller scale storage facilities.		
<b>‘Endangered’ in the Presence of Aggravating Conditions</b> Encryption; poor handling; poor storage; lack of consistent replication; failure of external (dependencies, e.g., suppliers, security); political or commercial interference; failure of internal dependencies (e.g., power supply, disk controller); overly aggressive compression; poor information security; lack of integrity-checking; lack of strategic investment; lack of warranty; unenforceable warranty; encryption.		
<b>‘Lower Risk’ in the Presence of Good Practice</b> Backup to different technology; backup to diverse locations; documentation of assets; integrity checking; preservation planning; refreshment planning; export functionality; resilient to hacking; selection and appraisal criteria; version control; resilient funding; technology watch; enforceable warranty; disaster planning.		
<b>2021 Review</b> This entry was added in 2019 to ensure that the range of media storage is properly assessed and presented. It was reviewed in 2020 with no noted change towards greater or reduced risk. While the 2021 Jury agreed with the current <i>Vulnerable</i> classification, they added a trend towards greater risk in light of the continued shift towards reliance on cloud storage with computers increasingly reducing hard disk for solid-state storage and commercial motivations for less support. As people increasingly select other storage methods such as cloud, they are less likely to maintain existing content on portable hard disks, which means the portable hard disks are more likely to be overlooked or ignored (e.g., left in drawers) rather than checked and refreshed.		
<b>Additional Comments</b> See also: Digital Preservation Coalition (2016) ‘Storage’ in Digital Preservation Handbook (Second Edition), online at <a href="https://dpconline.org/handbook/organizational-activities/storage">https://dpconline.org/handbook/organizational-activities/storage</a> .		

<b>Local Network Storage</b> Materials routinely copied or backed up to locally managed data storage facilities and able to be restored under institutional service arrangements.		
Group: Storage	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Unanimous Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Institutional or departmental network storage and institutional data centers based on technologies such as (NAS) Network Attached Storage, (RAID) Redundant Array of Independent Disks, (SAN) Storage Area Networks, JBOD (Just a bunch of disks), SPAN and related		
<b>'Endangered' in the Presence of Aggravating Conditions</b> Encryption; lack of routine maintenance; lack of storage replication; over-dependence on a single supplier, technology or technician; insufficient documentation; single point of failure; political or commercial interference; failure of dependencies (e.g., power supply, controller software); overly aggressive compression; poor information security; lack of integrity-checking; lack of strategic investment; lack of warranty; unenforceable warranty, encryption.		
<b>'Lower Risk' in the Presence of Good Practice</b> Backup to different technology; backup to diverse locations; documentation of assets; integrity checking; preservation planning; refreshment planning; export functionality; resilient to hacking; selection and appraisal criteria; version control; resilient funding; technology watch; enforceable warranty; disaster planning and documentation.		
<b>2021 Review</b> This entry was added in 2019 to ensure that the range of media storage is properly assessed and presented. The 2021 Jury agreed with the <i>Vulnerable</i> classification with no trend towards greater or reduced risk.		
<b>Additional Comments</b> See <i>also</i> : Digital Preservation Coalition (2016) 'Storage' in Digital Preservation Handbook (Second Edition), online at <a href="https://dpconline.org/handbook/organizational-activities/storage">https://dpconline.org/handbook/organizational-activities/storage</a> .		

<b>PDF/A</b> Materials contained within PDF/A; a format derived from the PDF 1.4 variant but with additional requirements that reduce external dependencies. Includes more recent variants (PDF/A2 and PDF/A3)		
Group: Formats	Trend in 2021:	
Added to List: 2017	<b>Trend towards reduced risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within five years, detailed assessment within three years	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Documents stored offline, or online in repositories or EDRMS, including reports, agenda, minutes, correspondence, contracts, essays, articles, or research papers.		
<b>'Endangered' in the Presence of Aggravating Conditions</b> Lack of conformance or validation; Lack of preservation commitment or planning; Lack of storage replication; Poor management of data protection or intellectual property rights; Political or commercial interference; loss of version control; use beyond the design capability of the standard; complacency; poor metadata; disassociation from context; loss of authenticity; encryption.		
<b>'Lower Risk' in the Presence of Good Practice</b> Backup and documentation of media assets; validation on creation; part of preservation plan; export functionality; licensing enables preservation; preservation capability in repository or EDRMS service; resilient to hacking; selection and appraisal criteria; version control; authenticity; resilient funding and recognition of value; technology watch		
<b>2021 Review</b> This entry was added in 2017. PDF/A has become more stable and reduces dependencies, and thus curtails preservation risks for certain types of content. The 2021 Jury agreed and added that there is also now PDF/A-4, moving the trend towards reduced risk. However, this alone is not sufficient to ensure preservation, and therefore users should be warned against complacency. PDF/A has sometimes been misunderstood or misrepresented as a generic solution to all digital preservation requirements. In the eyes of the BitList Jury members, it can only offer a preservation solution when embedded within a wider preservation infrastructure.		
<b>Additional Comments</b> Vulnerability also depends on if the PDF file conforms to the specific PDF/A standard or not. This is caused by a combination of 1) not conforming to the standard and 2) collection managers assuming that the file is resilient simply because it purports to be a PDF/A. This risk is less with the format and more with the understanding and experience in data management. Moreover, materials embedded in or attached to PDF/A-2 and PDF/A-3 may be at risk.  <i>See also:</i> Fanning, B (2017) <i>Preserving with PDF/A (Second Edition)</i> , DPC Technology Watch Report 17-01 online at <a href="http://doi.org/10.7207/twr17-01">http://doi.org/10.7207/twr17-01</a> .		





<b>Pension, Mortgage and Insurance Records</b> Records of transactions for long-lived financial products and services contracted between individuals and corporations. These records typically contain or depend on significant amounts of personal information and outlast the infrastructure on which they were created.		
Group: Sensitive Data	Trend in 2021:	
Added to List: 2017	<b>No Change</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Applications, correspondence and ancillary records relating to pensions, mortgages and insurances and other contracts of long duration. This includes corporate databases, email, web archives and EDRMS, and may require some coordination of paper, microfiche, born-digital and digitized records. These records often include the scope and duration of the contract as well as any agreed changes during the lifetime of the product. It may also include evidence of mis-selling or other sharp practice, which only becomes apparent after the fact. This entry pertains to corporate records rather than personal records.		
<b>'Endangered' in the Presence of Aggravating Conditions</b> Lack of corporate preservation planning; lack of preservation within the procurement of corporate systems; companies conflating backup with preservation; loss of integrity and authenticity; loss of context and connections to provide meaning; lack of preservation capability within agencies; lack of preservation voice at executive level; poor planning and roadmap for corporate infrastructure; proliferation of legacy systems; slapdash procurement or migration of new systems; mergers and acquisitions leading to confusion of corporate systems; lack of compliance, audit or accountability at operational levels; encryption.		
<b>'Lower Risk' in the Presence of Good Practice</b> Backup and documentation; use of open formats and open source software; considered data management planning; licencing that enables preservation; preservation capability in designated repository; resilient to hacking; selection and appraisal in place; authenticity and integrity of records managed; resilient funding and recognition at executive level; technology watch; regular preservation audits; accreditation and participation in the professional preservation community.		
<b>2021 Review</b> This entry was added in 2017 but was outside the competence of the judges to assess at that time. It was assessed in 2019 with additional expertise invited to the panel to support this assessment and reviewed again in 2020. The 2021 Jury agreed with that 2019 assessment and subsequent 2020 review, which classifies these digital materials as <i>Vulnerable</i> with no trend towards greater or reduced risk.		
<b>Additional Comments</b>		




The work and outputs of the EDRMS Preservation Taskforce, such as the EDRMS Preservation Toolkit, may be helpful for guidance in this context. See:

<https://www.dpconline.org/digipres/implement-digipres/edrms-preservation-toolkit>

<b>Published Research Papers</b> Completed research papers published in serials, monographs or theses which fall under specific collecting policies of research libraries or archives and are managed through dedicated repository infrastructures.		
Group: Research Outputs	Trend in 2021:	
Added to List: 2017	<b>Trend towards reduced risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within five years, detailed assessment within three years	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Scholarly E-Books, Electronic Journals, E-theses, Electronic Monographs.		
<b>'Endangered' in the Presence of Aggravating Conditions</b> Lack of documentation; lack of clarity with respect to intellectual property; embedded complex objects; unstable funding for repository; lack of strategic investment; complex external dependencies; lack of persistent identifiers; bespoke formats; lack of legal deposit mandate.		
<b>'Lower Risk' in the Presence of Good Practice</b> Strong documentation including intellectual property rights; clarity of preservation path and ensuing responsibilities; credible preservation plan; proven capacity of repository; legal deposit preservation copying; post-cancellation access service; persistent identifiers used consistently; non-proprietary formats used and validated; minimal or well managed external dependencies.		
<b>2021 Review</b> This entry was added in 2017 under 'Published research outputs,' though without reference to the capacity of the repository infrastructure. The 2019 Jury amended it to presume the existence of repository infrastructure and noted that the aggravating conditions (which introduce risks) and good practice enhancements (which reduce it) are most relevant to repository operations. While the 2020 Jury found no change in trend, the 2021 Jury agreed it should remain <i>Vulnerable</i> but discussed improvements and initiatives towards the preservation of research data and outputs, leading to the trend towards reduced risk.		
<b>Additional Comments</b> See also: <ul style="list-style-type: none"> <li>Beagrie, N (2013) <i>Preservation, Trust and Continuing Access for E-Journals</i>, DPC Technology Watch Report 13-04, online at: <a href="http://doi.org/10.7207/twr13-04">http://doi.org/10.7207/twr13-04</a></li> <li>Morrissey, S, and Kirchhoff, A (2014) <i>Preserving E-Books</i>, DPC Technology Watch Report 14-01, online at: <a href="http://doi.org/10.7207/twr14-01">http://doi.org/10.7207/twr14-01</a></li> <li>Resources and recent outputs from Public Knowledge Project (PKP) Preservation Network, which developed to digitally preserve Open Journal Systems (OJS) journals, online at: <a href="https://pkp.sfu.ca/pkp-pn/">https://pkp.sfu.ca/pkp-pn/</a></li> </ul>		

Recently Commissioned or Completed Media Art		
Media art currently displayed in a gallery or in the process of being displayed.		
Group: Media Art	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards greater risk	Previously: Vulnerable
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Media art recently acquired by galleries that utilize specific hardware and software in order to be accessed or exhibited.		
<b>‘Endangered’ in the Presence of Aggravating Conditions</b> Lack of documentation to enable maintenance; lack of clarity with respect to intellectual property; complex interdependencies on specific hardware, software or operating systems; lack of capacity in the gallery or workshop; lack of strategic investment; complex external dependencies		
<b>‘Lower Risk’ in the Presence of Good Practice</b> Strong documentation; clarity of preservation path and ensuing responsibilities; proven preservation plan; capacity of workshop to support artwork at de-installation; capacity of gallery to conserve after de-installation; capacity of gallery to re-install work		
<b>2021 Review</b> This entry was added in 2019 as a separate entry, but it was previously introduced in 2017 under ‘Media Art’ with particular reference to historical media art. It was added for greater specificity for its recommendations, to represent works commissioned in the last five years where there is a reasonable expectation that documentation has been produced or could still be obtained. While the 2020 Jury found no change in trend, the 2021 Jury discussed how prospects for long-term preservation depend entirely on whether the artwork is collected post-commission and by an organization with the resources to care for it. The imminence of action is time-sensitive, requiring working with the artist to get the documentation from them about their work and what is needed before it is too late. Furthermore, there remains a vulnerability for the smaller museums or others that do not take the preservation of media art as seriously. For these reasons, the Jury agreed that the classification remains <i>Vulnerable</i> but with a trend changed towards greater risk.		
<b>Additional Comments</b> By the time digital art, time-based media, etc., has entered into the permanent care of a stewarding institution, many of its technologies are already end-of-life, unsupported, or the hardware components have deteriorated. Often the expertise to maintain these many interacting components sits outside the host organization, with a technical supplier to the gallery, and this is in itself vulnerable to business change. Although there are a few exceptions, there is a need for greater capacity within the museum and gallery sector to address the challenges.		

There have been new initiatives for guidance and examples of institutions taking wider sectoral responsibility for standards, which have helped with the effort to preserve, such as Matters in Media Art information resource and guidance. See: <http://mattersinmediaart.org/>


<b>Research Data Published through Repositories</b> Research data published through digital repositories or other services providers with specialist skills to manage the data and an ongoing commitment to ensure preservation.		
Group: Research Outputs	Trend in 2021:	
Added to List: 2019	<b>Trend towards reduced risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Recognized data repositories in specialist disciplines; institutional data repositories in subject specialist centres and partnerships		
<b>'Endangered' in the Presence of Aggravating Conditions</b> Lack of long-term commitment; lack of user community; lack of visibility to potential depositors; lack of institutional commitment		
<b>'Lower Risk' in the Presence of Good Practice</b> Certification and documented good practice; effective documentation requirements for depositors; proven financial sustainability; skilled staff; participation in the digital preservation community		
<b>2021 Review</b> This entry was added in 2019 as a separate entry, but it was previously introduced in 2017 under 'Research Data,' though without explicit reference to the capacity of the repository infrastructure. In 2019, the Jury split the 'Research Data' entry into a range of contexts for research outputs, including this addition. It was classified as <i>Vulnerable</i> ; the preservation of research data published through a well-founded repository with the capacity and commitment to ensure preservation and capability through their own professional development activities makes it a 'lower risk' outcome for research data. The 2021 Jury agreed with this classification but commented on the improvements and initiatives towards the preservation of research data and outputs, leading to a trend towards reduced risk.		
<b>Additional Comments</b> See also: L'Hours, Hervé, Kleemola, Mari, von Stein, Ilona, van Horik, René, Herterich, Patricia, Davidson, Joy, Rouchon, Olivier, Mokrane, Mustapha, & Huber, Robert. (2021). <i>FAIR + Time: Preservation for a Designated Community</i> (01.00), online at: <a href="https://doi.org/10.5281/zenodo.4783116">https://doi.org/10.5281/zenodo.4783116</a>		

## 9. Endangered



Digital materials are listed *Endangered* 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 *Vulnerable* materials in the presence of aggravating conditions, and *Critically Endangered* materials in the presence of good practice.


<b>3D Digital Engineering Drawings</b> 3D digital engineering models produced as part of building or engineering design processes. The production of such drawings has progressed from a digital analogue of paper to complex digital environments such as BIM (Building Information Modelling) which combine original drawings, libraries of compound objects, and links to external data sets such as about the performance of materials and maintenance of parts.		
Group: Engineering	<b>Trend in 2021:</b>	
Added to List: 2017	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Building Information Management (BIM), Computer Aided Design (CAD), Product Data Management in engineering and architecture,		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Lack of preservation mandate or collecting institution; lack of preservation capability in data owner; irregularities in supply chains; complex or long data supply chains; dependencies on proprietary software or formats; lack of persistent identifiers; poorly managed IPR; temporary joint-venture companies; poor records management; poor regulatory compliance; encryption.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Well managed data infrastructure; preservation from the point creation; carefully managed IPR; persistent identifiers; well managed records management processes; recognition of preservation requirements at highest levels; strategic investment in digital preservation; host clearly identified; participation in the digital preservation community.		
<b>2021 Review</b> This entry was first submitted in 2017 when the Jury lacked the capacity to consider it in detail. In 2019 it was assessed with additional expertise co-opted, with the decision to remain a very broad category, including major one-off construction and engineering projects, a long tail of more minor building programmes, and large volume but homogenous production processes in engineering. Both the 2020 and 2021 Jury agreed with its <i>Endangered</i> status. The key consideration is that the lifecycle of the products and the data that describes them vastly exceeds the short lifecycles of the infrastructures on which they are designed. This challenge is compounded by supply chains that may involve many different stages of production, as well as the delivery of large projects through transitory joint ventures companies that have no residual mechanism or capacity to preserve the data thereafter. Although there have been advancements in the development of new preservation tools and techniques for these materials, there are recent examples of the loss of 3D architectural drawings; these have had a huge impact, especially at the local level, as well as significant impacts on infrastructure, travel, and how people interact with built environments throughout the world. For these reasons, the trend has moved towards greater risk.		
<b>Additional Comments</b>		

Data in this category enables the safety and security of critical infrastructure, but the responsibility to maintain data is unclear, nor are retention periods clear. Although examples of good practice exist, the extent to which there are working solutions at large seems doubtful, and it is surprising that there are not more diverse success stories to report.

*Case Studies or Examples:*

- The Grenfell Tower Inquiry offers a case to consider how the loss of 3D Digital Engineering Drawings can have a huge impact, especially at the local level. For example, if Grenfell had been done using 3D technologies, do we have confidence that those materials would have been adequately preserved? What would have been the local impact? What would have been the impact on the inquiry? See: <https://www.grenfelltowerinquiry.org.uk/>




<b>Born Digital Photographs and Video shared via Social Media or Uploaded to Cloud Services</b>		
Digital images or video with no analogue equivalent and where the only copy is online with a social media platform or cloud image hosting service.		
Group: Social Media	<b>Trend in 2021:</b>	Consensus Decision
Added to List: 2018	<b>Trend towards greater risk</b>	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to address losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Flickr; Vimeo; YouTube; Instagram; Periscope; DropBox; Facebook; Twitter.		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> lack of preservation capacity in provider; lack of explicit preservation commitment or incentive from provider to preserve; lack of storage replication by provider; dependence on proprietary products or formats; poor management of data protection; inaccessibility to automated web crawlers; political or commercial interference; lack of offline equivalent; over-abundance; poorly managed intellectual property rights; lossy compression applied in upload scripts.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Offline backup; lossless compression; good documentation; access to web harvesting; clarity of intellectual property rights that enable preservation; credible preservation commitment from service provider; export pathway.		
<b>2021 Review</b> This entry was added in 2018 within a wider social media group, sharing common risk profiles and challenges with other social media materials such as dependency on global service providers whose business model can only be presumed and tied to users via asymmetrical contracts that favour the supplier. In 2019 it became a standalone entry given the distinct preservation challenges of images and videos in cloud services which need to be addressed. The 2020 Jury noted a trend towards increased risk based on cloud services with a low barrier to entry leading to use by agencies or individuals least able to respond to closure or loss.  While there have been developments and wider use of tools like pywb tools (e.g., the Webrecorder toolset) and API harvesting tools (twarc, TAGS), the 2021 Jury agreed with this continued trend towards increased risk, adding in light of recent global crises (predominantly the coronavirus pandemic, compounded by vaccine hesitancy, but also the increasing deterioration of the world's democracies) as a result of widespread misinformation, this category is now more significant and has a greater impact		
<b>Additional Comments</b> The vast majority of content may be accessible for as long as the platform where it is hosted is popular (and has a viable business model); however, more insidious content (such as malicious		

misinformation or hate speech) may be deleted by content creators (potentially backed by hostile governments) to avoid prosecution or tracing. It is unclear to what extent these platform providers are compelled to provide access to servers / deleted content or private content for evidential purposes in the course of legal or criminal investigations. The lack of transparency and standardized international regulation of these platforms make their content vulnerable to exploitation and malicious use by individuals, corporations, and hostile governments.

Museums, Libraries, and Archives have begun to pay attention to this content through projects like Collecting Social Photo (CoSoPho), but no breakthroughs have been made. See: <https://www.collectingsocialphoto.org/en/home>

*Case Studies or Examples:*


- Example of loss of users' content when social media platforms such as MySpace. See: Hern, A., (2019) Myspace loses all content uploaded before 2016. *The Guardian*. 18 Mar, online at <https://www.theguardian.com/technology/2019/mar/18/myspace-loses-all-content-uploaded-before-2016>.
- Examples of the short turnaround of closures within the photo-sharing community and point to the volatility in this market, such as the closure of the Ever photo storage service and changes to the Flickr free service. See: Perez, S., (2019) Ever, once accused of building facial recognition tech using customer data, shuts down consumer app. *TechCrunch*, online at <https://social.techcrunch.com/2020/08/24/ever-once-accused-of-building-facial-recognition-tech-using-customer-data-shuts-down-consumer-app>; and Gartenberg, C., (2018) *Flickr will end 1TB of free storage and limit free users to 1,000 photos*. The Verge, online at <https://www.theverge.com/2018/11/1/18051950/flickr-1000-photo-limit-free-accounts-changes-pro-subscription-smugmug>.


Cloud-based Services and Communications Platforms		
Digital content produced, stored and accessed within commercial cloud-based services and communications platforms. This entry broadly includes services based on a costed subscription and contract and also free online utilities offered at no cost to end-users, but with a business model based on gathering and reselling consumer insights.		
Group: Social media	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards greater risk	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Consumer Cloud-based Utilities: Google Docs, Google Sheets, Gmail, Hotmail, Yahoo Mail, Prezi Premium or institutional social media services: Premium versions of Vimeo, Flickr, Yammer, Slack, Microsoft Teams, and others.		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Unstable business model from service providers; abandonment of the service due to various reasons (e.g., service provider bought over or pivots to new market opportunities); lack of export functionality; unstable terms and conditions; lack of onsite copy of key media; lack of investment in infrastructure; lack of strategic plan for IT provision; confusion on IPR; conflating preservation and access.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Clear export and migration pathways; preservation responsibility shouldered by the service provider; Offline back up for key media; fit to preservation and records management plan; strategic roadmap for adoption of social media;		
<b>2021 Review</b> This 2021 entry merges two separate 2019 entries, ‘Consumer Cloud-based Utilities’ and ‘Premium or institutional social media.’ The 2021 Jury decision to merge the entries emphasizes the similarities and common threats faced by cloud services that are both 'paid-for' and 'free-at-the-point-of-use.' While there are differences, both share similar aggravating conditions relating to their dependencies on the vendor's business model and the terms and conditions they impose.		
<b>Additional Comments</b> While there are largely shared challenges between the merged entries, it should be noted that with digital materials from Consumer Cloud-based Utilities, the business model and sustainability can only be presumed, and contracts tend to be asymmetrical in favour of the supplier. Moreover, because these services have a low barrier to entry, they may be favoured by agencies or individuals least able to respond to closure or loss. If referring to the entire platforms and risk of the entirety of data on these, the concern is that the corporation providing the service suddenly decides it is no longer of value to them. In these circumstances, materials could be removed quickly. That has happened previously and will certainly be seen again. Preservation is not a commitment that most providers make.		

Similarly, with digital materials from premium or institutional social media services, the business model and sustainability are more obvious, and contracts may be enforceable more readily. Moreover, because these services have a slightly higher barrier to entry, they may be favoured by agencies better able to respond to closure or loss. Traditional web archiving can be employed where the user pays for a service, but the content is ultimately publicly available (such as Flickr). But much is unclear about how to preserve internal social media / closed networks that web archiving cannot get to or existing tools do not cover. The growth in the use of these products for communication and social networking led the Jury to choose the *Endangered* classification. Existing tools could be modified to tackle some of the closed networks. Still, it is likely to require investments, perhaps related to corporate records in some cases (thinking about internal Slacks, for instance), and more education about the importance of preserving this material and not trusting the publishing platforms to host the content forever.

*Case Studies or Examples:*


- The website 'Killed by Google' provides a list of projects and apps that Google has shut down over the years, dating back to 2006. See: <https://killedbygoogle.com/>

Completed Investigations based on Open Source Intelligence Sources		
Open source social media and web content that has been used to support the conclusions of crowd-sourced investigation and fact-checking in political or military conflict.		
Group: Digital Legal Records	Trend in 2021:	Unanimous Decision
Added to List: 2019	Trend towards greater risk	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Social media sources relating to recent conflicts, such as in Ukraine.		
<b>‘Critically Endangered’ in the presence of Aggravating Conditions</b> Encryption; loss of authenticity; lack of preservation agency; limited or no digital preservation capability.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Offline backup captured by a journalist or investigating authority; materials presented and documented in court; court able to deliver preservation; authenticity protected		
<b>2021 Review</b> This entry was added in 2019 and subsequently split into three elements by the Jury, relating to current, recent and historic sources. This entry relates to materials used in evidence in completed investigations, as well as those presented to courts or other investigatory agencies. Social media companies have a policy to take down or suppress content that they consider propaganda for terrorist groups. This has had the unintended consequence of deleting or suppressing content used in open source investigation or fact-checking for journalistic or judicial purposes, which may impede refutation or prosecution. However, a new generation of cloud-based services now allows investigators to copy and stabilize content to private accounts in the process of investigating it, so the ethical requirements of social media companies and the integrity of the investigation are both served. The 2020 Jury noted that such content remains at risk. The presentation of data to a court or prosecuting authority, or the publication through news media, implies the introduction of a long-term preservation function. The 2021 Jury agrees with this assessment and <i>Endangered</i> classification but has changed the trend towards greater risk in light of recent developments in crowd-sourced investigations and fact-checking		
<b>Additional Comments</b> See: Higgins, E. (2019). <i>Bellingcat and beyond. The future for Bellingcat and online open source investigation</i> , iPres Conference 2019, Amsterdam, <a href="https://www.youtube.com/watch?v=kZAb7CVGmXM">https://www.youtube.com/watch?v=kZAb7CVGmXM</a> .		

Content on Cloud Video Services Produced by the Service Provider		
Video materials - films and television programs – that are produced by companies that maintain their own distribution platforms and are exclusively available through these platforms.		
Group: Sound and Vision	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards greater risk	Previously: Vulnerable
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Netflix, Amazon Prime		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Lack of corporate preservation planning; lack of preservation capability; conflating backup with preservation; loss of original recordings; lack of preservation voice at executive level; poor planning and roadmap for infrastructure; slapdash procurement or migration to new systems; mergers and acquisitions; profusion of corporate systems; complex intellectual property rights; single point of failure; technical protection measures that inhibit reasonable preservation actions.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> backup and documentation; use of open formats and open source software; data management planning; licencing that enables preservation; corporate preservation capability; resilient to hacking; authenticity and integrity managed; recognition of preservation functions at executive level; technology watch; preservation audits; participation in the preservation community.		
<b>2021 Review</b> This entry was added in 2019 to represent collections that are highly significant in cultural and social terms. It was adopted as the Jury was unclear whether the content could be played outside of the producers' publication platform. This introduces technical dependencies between content and software, which are amplified by rights management. The 2021 Jury agrees with the <i>Endangered</i> classification but discussed how the growth of content produced with no or limited preservation mandate has led to greater risk.		
<b>Additional Jury Comments</b> This entry has five aspects. 1. It falls outside the scope of traditional regulatory frameworks and archiving has not yet been included in any legislative framework, unlike broadcast TV, where there is a designated archive in most developed nations. 2. As a result, the collecting and preservation of the content from online platforms is underdeveloped, and the content remains unavailable in public archives. 3. These risks are mitigated by the fact that the commercial archives are technologically advanced, with mature digital ecosystems and skills, and much of the content has a 'long tail' business model, and as commercial products have value, so preservation incentives are clear. 4. However, these are often stored at scale on LTO tapes, and so specific		


issues arise with the obsolescence of LTO tape technologies for the broadcast sector. 5. Nonetheless, issues remain around archiving relevant assets which may not be valued by the production company.




<b>Contractual Documents and Related Records</b> Documents, correspondence and other records created in the course of contractual dealings between individuals and agencies, especially where the subjects are of long duration and may be subject to legal scrutiny at undefined points in the distant future.		
Group: Digital Legal Records	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Contracts, receipts, correspondence, license agreements, building consent, warranties, and any other document or record that represents a legally binding transaction or permission. Such records may be useful in the avoidance or resolution of disputes, whether in court or prior to proceedings. Includes Online Terms and Conditions for e-commerce or end-user agreements for services.		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Loss of context; loss of authenticity or integrity; external dependencies; poor storage; lack of understanding; churn of staff; poorly framed or over-zealous disposal; ill-informed records management; misplaced fears with respect to data protection, encryption.		
<b>'Vulnerable' in the Presence of Good Practice</b> Well managed data infrastructure; preservation enabled at the point of creation; carefully managed authenticity; use of persistent identifiers; finding aids; well managed records management processes; application of records management standards		
<b>2021 Review</b> This entry was added in 2019 as a subset of an entry introduced in 2017 for 'Digital Legal Records and Evidence,' which was split into four more discrete entries. There is overlap with Pension Mortgage and Insurance Records entry, but this entry addresses digital records with value over the long term that may not be immediately obvious at the point of creation. The likelihood of liquidation, mergers or acquisitions means that these records are trending towards greater risk. There was a 2020 trend towards greater risk in light of the Covid Pandemic causing profound dislocation across the economy, placing many companies and agencies at financial risk. The 2021 Jury, however, noted some improvements in record-keeping legislation and regulations and guidance to this area advising what needs to be kept and for how long; while there is still room for improving how legal agreements are managed through records management standards and processes, there is a growing number of agencies with experience of handling this and therefore there has not been a notable trend towards an increased risk for 2021.		
<b>Additional Comments</b>		





The work and outputs of the EDRMS Preservation Taskforce, such as the EDRMS Preservation Toolkit, may be helpful for guidance as many of these records will be held in EDRMS type systems  
<https://www.dpconline.org/digipres/implement-digipres/edrms-preservation-toolkit>.


Corporate Records of Long Duration on Network Drives, Intranets and EDRMS		
Records on internal corporate network drives, intranets or document management services where access is limited to a distinct group of users, and in which the lifecycle of the record or the business processes they support is greater than the technology on which they are created or retained.		
Group: Sensitive Data	Trend in 2021:	Consensus Decision
Added to List: 2017	Trend towards greater risk	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to preserve materials in this group, with the development of new preservation tools or techniques.
<b>Examples</b> Born-digital records of small and medium sized enterprises; fast-changing internal manuals, advice or policies shared on intranets or EDRMS; records of long-lived products and services; Historic guidelines and manuals which evidence 'best practice'; Documentation supporting long-lived contractual relations; Online terms and conditions; Corporate Slack channels; Google Drives; EDRMS; Email.		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Lack of systematic preservation function; lack of preservation path or plan for data; dependence on proprietary products or formats; poor management of data protection; political or commercial interference; lack of offline equivalent; over-abundance through poor disposal or version control; lack of capacity; lack of commitment; loss or lack of documentation; sector-specific software or data types; encryption		
<b>'Vulnerable' in the Presence of Good Practice</b> Preservation infrastructure and pathways; replication; appraisal and selection including de-duplication;		
<b>2021 Review</b> This entry was added in 2017 to draw attention to the pressing need for digital preservation in business, especially in small to medium enterprises. The 2020 Jury noted how the Covid Pandemic has caused profound dislocation across the economy and placed many companies and agencies at financial risk. The likelihood of liquidation, mergers or acquisitions means that these records are trending towards greater risk. The 2021 Jury agrees with the trend towards greater, adding that increased risk is not necessarily because there are no assigned parent archives to take on these materials; rather, it is because they too often sit in these spaces for some time before being transferred to the archives. They are often not well managed or maintained by their creating agencies, putting them at risk of accidental deletion or corruption. There remain increased risks without business continuity and trust.		
<b>Additional Comments</b> Corporate records should form part of organizational records management schemes, and so responsibilities should be clear; however, this may be much more challenging for smaller organizations without dedicated roles or with complex data types.		

The work and outputs of the EDRMS Preservation Taskforce over the last year, for example the EDRMS Preservation Toolkit, may be helpful for guidance and examples in this context  
<https://www.dpconline.org/digipres/implement-digipres/edrms-preservation-toolkit>.

<b>Current Portable Magnetic Media</b> Materials saved to magnetic tape, portable hard disks or other magnetic media in the last five years where the reader devices are still supported and can be integrated easily into hardware infrastructure.		
Group: Portable Media	Trend in 2021:	
Added to List: 2019	<b>Trend towards reduced risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> LTO tapes; portable hard disks		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Poor storage conditions; encryption; digital rights management; lack of replication; lack of documentation; lack of periodic testing; lack of refreshment pathway; lack of access to readers; out of manufacturers' warranty or no warranty; storage within paper files.		
<b>'Vulnerable' in the Presence of Good Practice</b> Regular review and testing; replication; refreshment plan; comprehensive documentation; high quality storage; regular maintenance of readers; multiple readers available.		
<b>2021 Review</b> This entry was added in 2019 to ensure that the range of media storage is properly assessed and presented. Magnetic media is typically more fragile than optical media because it is susceptible to 'bitrot' and magnetic damage in ways that optical media are not. However, the 2021 Jury commented on how the types of magnetic media used have improved the last five years, notably the use of LTO, increasing good practice and trending towards reduced risk in this respect. It is important to note that LTO tapes come in different generations. Some pose greater preservation risks now (e.g., an organization with no equipment or way of reading content), so the use of LTO is good practice so long as it includes the active management of associated risks.		
<b>Additional Comments</b> This entry is highly dependent on who is looking after the portable media but made more difficult over time. The lack of granularity in the definition means that only general advice can be offered, such as to refresh media. In time, it may yet be more useful to split all storage media (maybe 100 items long) with an indication of how long these can be expected to last. In many cases, specialists can recover obsolete media, but the cost of employing them can become an aggravating condition. It is important to emphasize that the short lifetime of many storage devices is not a problem to be solved with new long-lasting storage technologies (and indeed, many inventions have come and gone). Cheap commodity storage has been purposely designed to deliver value at a low price for a short time. Therefore, management and preservation processes for monitoring and refreshment need to take these characteristics into account.		


<b>Current Portable Optical Media</b>		
Materials saved to DVDs, CDs or other optical media in the last five years where the reader devices are still supported and can be integrated easily into hardware infrastructure		
Group: Portable Media	Trend in 2021:	Unanimous Decision
Added to List: 2019	<b>No Change</b>	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> CDs, DVDs produced in the last five years		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Poor storage conditions; encryption; digital rights management; lack of replication; lack of documentation; lack of periodic testing; lack of refreshment pathway; lack of access to readers; out of manufacturers’ warranty or no warranty; storage in paper files.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Regular review and testing; replication; refreshment plan; comprehensive documentation; high quality storage; regular maintenance of readers; multiple readers available;		
<b>2021 Review</b> The judges have introduced this entry to ensure that the range of media storage is properly assessed and presented.		
<b>Additional Comments</b> This entry is highly dependent on who is looking after the portable media but made more difficult over time. The lack of granularity in the definition means that only general advice can be offered, such as to refresh media. In time, it may yet be more useful to split all storage media (maybe 100 items long) with an indication of how long these can be expected to last. In many cases, specialists can recover obsolete media, but the cost of employing them can become an aggravating condition.		

<b>Current Portable Solid-State Media</b> Materials saved to flash drives or other solid-state media in the last five years where the reader devices are still supported and can be integrated easily into hardware infrastructure		
Group: Portable Media	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Unanimous Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> USB memory sticks; flash cards in cameras; solid state portable hard disks		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Poor storage conditions; encryption; digital rights management; lack of replication; lack of documentation; lack of periodic testing; lack of refreshment pathway; lack of access to readers; out of manufacturers' warranty or no warranty; storage in paper files.		
<b>'Vulnerable' in the Presence of Good Practice</b> Regular review and testing; replication; refreshment plan; comprehensive documentation; high quality storage; regular maintenance of readers; multiple readers available;		
<b>2021 Review</b> This entry was added in 2019 to ensure that the range of media storage is properly assessed and presented.		
<b>Additional Comments</b> This entry is highly dependent on who is looking after the portable media but made more difficult over time. The lack of granularity in the definition means that only general advice can be offered, such as to refresh media. In time, it may yet be more useful to split all storage media (maybe 100 items long) with an indication of how long these can be expected to last. In many cases, specialists can recover obsolete media, but the cost of employing them can become an aggravating condition.		

Digital Music and Ephemera Shared on Social Media		
Digital materials created by musicians and fans as a by-product of performance or recording, shared on websites and other social media platforms.		
Group: Sound and Vision	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards greater risk	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to address losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Fan sites; private or illicit recordings of concerts; informal music sharing between networks such as TikTok, MySpace and Facebook		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Dependence on social media provider; lack of offline equivalent; infringing intellectual property right; unstable or small community of interest; encryption.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Offline equivalent; intellectual property rights conducive to preservation; partnership with collecting institution; availability to web archiving.		
<b>2021 Review</b> In 2019, this entry was created as a subset of a previous 2017 entry, 'Digital Music Production and Sharing,' which was split to draw attention to the different challenges faced by the different forms. This entry overlaps with other entries relating to social media as well as those relating to community-generated content but is a separate entry to emphasize the context in which music is shared and enjoyed. This context could be lost if our attention were on products controlled by studios or artists. The 2021 Jury discussed content increasingly being shared across multiple platforms, which is both good and bad for risk. A multi-platform nature provides an element of protection against total loss, but the role and type of interaction with the content on each platform are also important and expanding with limited attempts at preservation. For these reasons, the trend has moved towards greater risk with the need for selective approaches given the increasing volume of material.		
<b>Additional Comments</b> The ephemera are increasingly stored on websites that themselves are fragile and are removed, and nothing held on these services can be relied on in archival timeframes.  Web archiving and social media archiving have matured, so a representative sample is probably readily available for particular countries which are more mature in their digital preservation activities as opposed to other countries which are not.		
Case Studies or Examples:		


- The recent case of 'Yahoo Groups' closure serves to underline the fragility of community content hosted by third parties. See Brinkmann, M. (2020). *Farewell Yahoo Groups! Shutting down on December 15, 2020*. gHacks Technology News, <https://www.ghacks.net/2020/10/14/farewell-yahoo-groups-shutting-down-on-december-15-2020/>
- The case of MySpace's removal of MP3s demonstrates a big loss of shared digital recordings through the platform and subsequent recovery efforts by an academic group and Internet Archive. See Kleinman, Z. (2019). *MySpace admits losing 12 years' worth of music uploads* - BBC News, <https://www.bbc.co.uk/news/technology-47610936> and Internet Archive. *The Myspace Dragon Hoard (2008-2010)*, [https://archive.org/details/myspace\\_dragon\\_hoard\\_2010](https://archive.org/details/myspace_dragon_hoard_2010).
- Flickr's decision to delete vast numbers of images held on free accounts and material lost as a result of Tumblr's policy changes on the definition of adult content. See: Tiffany, K. (2019). *Flickr will soon start deleting photos — and massive chunks of internet history*, Vox, <https://www.vox.com/the-goods/2019/2/6/18214046/flickr-free-storage-ends-digital-photo-archive-history>



<b>Digital Radio Recordings</b> Master recordings of radio broadcasts generated live but often poorly stored thereafter, for example offline recordings on single LTO (Linear Tape Open) Tapes		
Group: Sound and Vision	Trend in 2021:	
Added to List: 2017	<b>No Change</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Broadcast archives of UK commercial local radio;		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Lack of archival mandate; lack of capability of archive; lack of policy or capacity within broadcaster; small or unprofitable broadcaster; concern over intellectual property rights; overzealous rights management protection; device or software dependence; dependence on proprietary or obsolete formats; lack or loss of documentation; little use or inaccessibility; storage (typically tapes) older than warranty; lack of media refreshment plan; lack of error or integrity checking process; single copies		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Archival responsibility accepted and acted upon; replication; refreshment of media; good documentation; active in digital preservation community; trusted repository; content re-used		
<b>2021 Review</b> This entry was added in 2017 as a separate entry due to concern over recordings on LTO tapes. These provide between 15 and 30 years’ storage which may be less depending on usage and storage conditions. LTO1 and LTO2, which were released in 2000 and 2003 respectively, have largely reached the final phases of viability. Reader compatibility may be more problematic than media resilience, however. Drives supporting newer releases of the format are typically only compatible within two generations, and experience with the recently released LTO8 suggests that it is only backwardly compatible to one generation. Therefore one major national archive and library has decided to expedite migration away from LTO6, which is becoming obsolete more quickly than anticipated. Through time, the risks to collections that have not been refreshed or replicated from early LTO tapes expand. Thus, the overall trend is towards greater risk when collections are not migrated. Older formats, perhaps as recently as LTO6, extinction events should be anticipated within two to five years. The 2021 Jury agreed with the <i>Endangered</i> classification, noting the importance of a selective approach.		
<b>Additional Comments</b> Depending on the legislative context, digital radio recordings may fall under published works if they are broadcasts. Additionally, if the recordings are broadcast by a taxpayer-funded broadcaster, record-keeping guidelines may already exist to advise how long content be kept,		


which would then inform a selective approach to caring and looking after them for as long as required.


While broadcasters may keep their own programmes, they are often not comprehensively collected or archived by memory institutions as contributions to cultural heritage.

Digital Recordings Published via Cloud-based Music Sharing Platforms		
Music licensed and playable through corporate platforms protected by rights management and subscription revenues and presented as compressed single-track recordings.		
Group: Sound and Vision	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards greater risk	Previously: Vulnerable
Imminence of Action Action is recommended within three years, detailed assessment within one year.	Significance of Loss The loss of tools, data or services within this group would impact on people and sectors around the world.	Effort to Preserve It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
Examples Spotify, iTunes, Bandcamp, SoundCloud		
'Critically Endangered' in the Presence of Aggravating Conditions Lack of preservation capability within corporate systems; conflating backup with preservation; loss of original multi-track recordings; lack of preservation voice at executive level; poor planning and roadmap for corporate infrastructure; slapdash procurement or migration to new systems; mergers and acquisitions; profusion of corporate systems; complex intellectual property rights; single point of failure; technical protection measures that inhibit preservation actions; encryption.		
'Vulnerable' in the Presence of Good Practice Strong backup and documentation; use of open formats and open source software; data management planning for preservation; licencing that enables preservation; corporate preservation capability; resilient to hacking; authenticity and integrity managed; recognition of preservation functions at executive level; technology watch; regular preservation audits; accreditation and participation in the professional preservation community.		
2021 Review This entry was previously under the 2017 'Digital Music Production and Sharing,' split by the 2019 Jury into four subsets, recognizing the different challenges faced. This entry is particularly concerned with the music industry at scale and the services that connect the vast majority of artists to their audiences. These are typically large and well-funded, and typically recognize the value of the content they publish. But this is not without risks. It is perhaps surprising that the music industry does not yet have any equivalent to the non-print legal deposit regime that applies to other types of publication, including sheet music in some jurisdictions. The 2021 Jury noted a large amount of vulnerable material on user-driven platforms where material can be very ephemeral (removals resulting from, e.g., account deletion, space limitations, copyright claim), and the issue of license and licensing with the instability of the business model. For this reason, the scope is widened to include ad hoc sharing so that this entry broadly includes all platforms such as SoundCloud, Bandcamp which are more community-driven, as well as Spotify, resulting in a raised classification and trend towards greater risk.		
Additional Jury Comments		


The preservation of recorded music is one of our generation's most important jobs, but it is unclear where responsibility lies. There are commercial incentives to do so, but there are also incentives to reduce costs. Whilst public archives are permitted to keep this material in some jurisdictions they typically do not have the resources to do so. Consequently, there is an expectation that rights holders will maintain their own archival copies but may not do so. National collecting organizations may need to develop a role to address this.

If managed well, there is hope. It may not be an issue in the cases where the production company would hold master recordings and, if a streaming service lost a track (e.g., Spotify), they would go to the production company and ask for a copy. However, it is an issue for those outside of production companies and platforms such as SoundCloud and Bandcamp, which are more community-driven.

<b>Digitally Published Sheet Music</b>		
Sheet music is licensed and published in various digital formats and subject to copyright restrictions and often protected by digital rights management technologies.		
Group: Sound and Vision	Trend in 2021:	Unanimous Decision
Added to List: 2019	<b>No Change</b>	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within five years, detailed assessment within three years.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a small effort to address losses in this group, requiring the application of proven preservation tools or techniques.
<b>Examples</b> This entry includes all manner of published sheet music, including choral works, orchestral works, scores published in different forms such as PDF, PDF/A and XML.		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Encryption; uncertainty over intellectual property rights; uncertain business model of publisher; lack of legal deposit mandate.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Rights management conducive to preservation; held in a trusted repository; legislation in place such as legal deposit enabling copying.		
<b>2021 Review</b> In 2019, this entry was added as a subset of a previous 2017 entry, 'Digital Music Production and Sharing,' which was split to draw attention to the different challenges faced by the different forms. This entry focuses on digitally published sheet music, which often takes the form of PDF and PDF/A and E-Book formats. The 2021 Jury made no changes to this entry classification or trend but did comment on the importance of legislation (such as legal deposit) to manage associated risks.		
<b>Additional Comments</b> The commercial value of these materials should be a protection against their loss, and the history of the music industry indicates that sheet music continues to have value, so even if an individual organization fails or its DRM servers go offline, and some music becomes inaccessible, it is not lost. However, the difficulties of archiving DRM-locked files remain real.		

<b>Electronic Hospital and Medical Records</b>		
Personal medical records and records of hospital treatment are increasingly—if not uniformly—born digital. By implication, those records should be retained through the lifetime of the patient, or in some instances longer as required for intergenerational study; and yet there is little evidence of the medical profession participating in the digital preservation community.		
Group: Sensitive Data	<b>Trend in 2021:</b>	Consensus Decision
Added to List: 2017	<b>Trend towards greater risk</b>	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Medical scans; records of treatment and care plans; health advice and notifications;		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Loss of context; loss of authenticity or integrity; poor storage; lack of understanding; churn of staff; significant volumes of data; significant diversity of data; ill-informed records management; poorly developed transfer and integrity checking; poorly developed migration or normalizations specifications; longstanding protocols or procedures that apply unsuitable paper processes to digital materials; encryption		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Well managed data infrastructure; preservation enabled at the point of creation; carefully managed authenticity; use of persistent identifiers; well-managed records management processes; application of records management standards; recognition of preservation requirements at highest levels; strategic investment in digital preservation; preservation roadmap; participation in the digital preservation community.		
<b>2021 Review</b> This entry was first submitted in 2017 under 'Medical and hospital records.' At that time, there was limited capacity to address the topic. It was published as 'of concern' to revisit and review by the 2019 Jury and also independently received as a submission to the open nomination process under 'Electronic hospital and medical records.' The entry covers a broad range of material, and it may be useful in future years to split the entry into more discrete entries. Still, the 2021 Jury agreed to keep the current description and classification to draw attention to the scale of the digital preservation challenges which arise in hospitals and the medical profession. Additionally, the same reasoning for greater risk in 2020 was used for 2021; there has been significant strain through the Covid pandemic, with resources stretched to meet an overwhelming demand and rigid, exacting protocols. In this environment, it is hard to avoid the sense that records are also now at greater risk. The 2021 Jury further commented that hospital records are at greater risk than we think, where there may already be poor maintenance of records during their lifecycle, poor migration planning, etc.		
<b>Additional Comments</b>		

Increasing sensitivity and awareness of data protection requirements could act inadvertently as a barrier to lifecycle data management. It is striking how little evidence is of the health technology companies participating in the global digital preservation community.

<b>Email</b> Documents, correspondence and other records created in the course of contractual dealings between individuals and agencies, especially where the subjects are of long duration and may be subject to legal scrutiny at undefined points in the distant future.		
Group: Formats	Trend in 2021:	
Added to List: 2017	<b>Trend towards reduced risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools or services within this group would have a global impact.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Email in all its forms, including individual messages, threads of conversation, mailboxes, email servers and file attachments.		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Conflicting and unmanaged IPR; use of personal accounts for professional work and vice versa; proliferation and duplication of attachments; email not recognized as a record; absent, unworkable or inconsistent records management; dependence on free cloud-based services; lack of migration path; lack of preservation planning; perverse incentives to delete; encryption.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Application of appraisal and selection tools; timely transfer to preservation facility or archive; commitment to transparency; preservation policy; working preservation plan; clear migration path; widespread recognition of email as a record.		
<b>2021 Review</b> This entry was added in 2017, but the Jury did not have the capacity to assess it in detail. It was reviewed and assessed in 2019, including highlights to significant developments, including the recommendations of the Email Preservation Taskforce and the development of the ePADD software. Email presents many preservation challenges, from scale through core technologies, attachments, privacy and intellectual property rights. Because this entry intersects with many others, the aggravating conditions associated with email should be considered in conjunction where relevant. The 2021 Jury discussed the continued developments in email preservation tools and techniques as well as the growing number of archives preserving email content. At the same time, issues with providing access to preserved email content have arisen. Ongoing records management policies towards corporate or business email need to be better embedded to stop the loss of important email content, and more awareness is needed around the potential of personal email. While record-keeping legislation and mandates direct retention periods, email document decisions taken by government officials at local, regional and national levels are not always well maintained, if at all; a loss could impact people's lives along with their ability to assert rights. For these reasons, there is a 2021 trend towards reduced risk, but the <i>Endangered</i> classification remains.		




**Additional Comments**

Email is hugely important as it has been so pervasive as a communication mechanism for society. Some methods used and responsibility adopted for collecting at the business and public body level (again will differ globally), but this will be a fraction of the communities that use it, and few will be set up for the long-term care of this data.


*See also:*

- Prom, C (2019) Preserving Email (2<sup>nd</sup> Edition) DPC Technology Watch Report online at: <http://doi.org/10.7207/twr19-01>
- Artefactual Systems and DPC (2021) Preserving Email DPC Guidance Notes online at <http://doi.org/10.7207/twgn21-08>
- Murray K and Prom, C (2018) The Future of Email Archives: A Report from the Task Force on Technical Approaches for Email Archives, CLIR, online at: <https://clir.wordpress.clir.org/wp-content/uploads/sites/6/2018/08/CLIR-pub175.pdf>


<b>Legacy Media Art</b> Media art in storage or not otherwise displayed but where the artists or technicians are available to support installation		
Group: Media Art	<b>Trend in 2021:</b>	
Added to List: 2019	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within 12 months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples:</b> Media art in storage		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Lack of documentation to enable maintenance; lack of clarity with respect to intellectual property; complex interdependencies on specific hardware, software or operating systems; lack of capacity in the gallery or workshop; lack of strategic investment; complex external dependencies; loss of institutional memory resulting from staff churn; poor working relationship between the gallery and artist/workshop; lack of conservation assessment.		
<b>'Vulnerable' in the Presence of Good Practice</b> Strong documentation; clarity of preservation path and ensuing responsibilities; proven preservation plan; capacity of workshop to support re-installation; capacity of gallery to conserve; capacity of gallery to re-install; retention of institutional memory including archives of correspondence between gallery and artist/workshop; strong and continuing working relationship between the gallery and artist/workshop; regular conservation assessment.		
<b>2021 Review:</b> Media Art was introduced in 2017, though with particular reference to historical media art. The 2019 Jury added and scoped this entry to ensure greater specificity in its recommendation. It is intended to represent works held in galleries but no longer displayed, but where there is a continuing working relationship between the gallery and the artist or workshop and reasonable expectation that support for preservation could still be obtained when required. The 2020 review found a trend towards greater risk, given that many museums and galleries, which often rely on visitors for income, have been closed for extended periods. Moreover, any form of digital materials that rely on an individual's knowledge is at particular risk by a pandemic. For similar reasons, the 2021 Jury agreed with the continued trend towards greater risk, noting that digital materials in museums and galleries records are likely to be at greater risk in these circumstances.		
<b>Additional Comments</b> This entry attempts to capture a point in the lifecycle of media art where preservation risks are increasing but not yet critical. There is a risk that preservation issues will not become apparent until the piece is brought out of storage when considered for loan or exhibition - often on timescales that make it too late to address preservation concerns effectively. Galleries should be		

aware that the range of data/formats/hardware/software embedded in media art can be wide and vary at different speeds.


Sooner action is needed to prevent the material from becoming *Critically Endangered* once the artist has died or relationships break down. Where the artist is still around, there is a major reduction in the inevitability of loss and its potential to be a potentially newsworthy subject. The loss of it would be just as impactful and significant though

<b>Legacy Video Files</b> Video files in any format containing moving pictures and sound recordings, particularly those that are proprietary, contain or utilize encrypted Digital Rights Management (DRM) or carrier-bound.		
Group: Sound and Vision	<b>Trend in 2021:</b>	
Added to List: 2019	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent losses in this group, such as the development of new preservation tools or techniques.
<b>Examples</b> AVI; MOV; MKV; MP3; MP4; on DVD or other carriers		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Lack of replication; encryption; digital rights management; proliferation of file formats; weak or non-existent technical documentation; lack of preservation capability or commitment; poorly managed or digitization processes or QA; reliance on encoding/decoding software		
<b>'Vulnerable' in the Presence of Good Practice</b> Effective replication; normalization of file formats; strong technical documentation; preservation pathway; good descriptive cataloguing; trusted repository.		
<b>2021 Review</b> This entry was added in 2019 under 'Video files' to emphasize the issues of video preservation that pertain to offline recording, whether from broadcast, film industry, institutional and private collections. The 2019 Jury noted the connections between this entry and others relating to social media but argued for a standalone to emphasize the range of issues tied to numerous formats and standards. The 2021 Jury discussed the need for further rescoping, arguing that the entry was too broad to be useful without specifying at risk types or formats. For this reason, its scope was narrowed to legacy videos that are proprietary, encrypted or carrier-bound. The 2021 classification remained <i>Endangered</i> with a 2021 trend towards greater risk given the growing content of at-risk legacy video files but a limited mandate.		
<b>Additional Comments</b> There are simply too many formats and too many standards, but the FFmpeg project and its related tools have significantly mitigated the technical risk to most video files. This enables a practitioner to transform the vast majority of file formats to safer preservation formats while retaining significant properties. However, technical risk is only one of the factors. There needs to be institutional engagement with audio-visual data as a priority. The issue then becomes one of identifying the organizations responsible and, constrained by the cost to store video data, making effective selection decisions.  <i>See also:</i> NFSa, 2015. <i>Deadline 2025: collections at risk</i> . Of note, on page 04, "Tape that is not digitized by 2025 will in most cases be lost forever as: Analogue video and audiotape, as well as early digital tape formats, will be effectively inaccessible due to the practical inability to maintain		

playback systems". Online at: <https://www.nfsa.gov.au/corporate-information/publications/deadline-2025>

<b>Master Digital Music and Sound Recordings</b> Master recordings of music and other performance from which retail products are derived, typically in multiple tracks and uncompressed high-resolution sound quality		
Group: Sound and Vision	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Unanimous Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a small effort to address losses in this group, requiring the application of proven preservation tools or techniques.
<b>Examples</b> Master recordings owned by music industry		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Single point of failure; storage on old or degrading media; lack of ongoing investment in changing preservation requirements; lack of capability; poor documentation; dependence on small staff		
<b>'Vulnerable' in the Presence of Good Practice</b> High quality storage; meticulous and consistent replication; trusted repository; preservation requirement understood at the executive level and funded accordingly; leadership in preservation community; expert staff		
<b>2021 Review</b> In 2019, this entry was added as a subset of a previous 2017 entry, 'Digital Music Production and Sharing,' which was split to draw attention to the different challenges faced by the different forms. Though it overlaps with other entries, including 'Digital Archives of Music Production,' it is a separate entry to emphasize the inherent and great value of master recordings over and above those distributed and the concomitant need for active preservation.		
<b>Additional Comments</b> The imminence of action will depend on format and age, and the significance of loss may be more largely felt if recordings of a major recording star  This is interesting as the recording houses should be seeing the value of these - so why are they not taking responsibility for looking after them? Do they not feel it is in their financial interests? The archival practices of the studios are typically based on value - the recordings are assumed to be worth keeping. However, this means relatively low-value masters may not be transferred to new media in a timely way and could be lost. There is no comprehensive deposit scheme to address the long tail of music production, and it is often unclear exactly where responsibility lies.  <i>Case Studies or Examples:</i> <ul style="list-style-type: none"> <li>The Jury was particularly aware of the case of the 'Universal Fire' of 2008, which has only recently been reported widely. See Rosen, J., (2019) The Day the Music Burned. <i>The New</i></li> </ul>		

*York Times*. 11 Jun. <https://www.nytimes.com/2019/06/11/magazine/universal-fire-master-recordings.html>


<b>Orphaned Works</b> Digital materials where copyright is uncertain, disputed or unknowable, meaning that preservation actions are constrained or prevented.		
Group: Orphaned Works	<b>Trend in 2021:</b>	
Added to List: 2017	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to address losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Photographs, music recordings, literature.		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Lack of documentation; dependencies resulting from hardware, software or media; lack of use resulting in lack of priority; lack of strategic investment in digital preservation; workflows that inhibit preservation of content that has not been licensed; encryption; poor storage.		
<b>'Vulnerable' in the Presence of Good Practice</b> Preservation pathway enabled; proven preservation plan applied; active effort to resolve IPR issues; institutional willingness to take risks for preservation.		
<b>2021 Review</b> When this entry was added in 2017, there was little evidence of any renewed effort to address the issue of orphaned work. While there have been improvements to the baseline competence of the archival and library professions in their understanding of copyright and the skills to preserve contents, this alone provides a narrow basis for optimism and the scale of the challenge is likely to have grown just as quickly if not more so as aggravating conditions become more prevalent too. The 2021 Jury added that while content is preservable, the preservation of orphan works is a matter of process and risk appetite. Added to the complexity are changes to copyright legislation in and across different national and regional contexts, particularly for UK institutions post-Brexit, as noted in the additional comments below. For this reason, there is a 2021 trend towards greater risk.		
<b>Additional Comments</b> The Jury would encourage organizations to take a risk-based approach which would help them preserve collections. Copyright infringements are only likely to become a significant issue in the context of access, and in most cases, the likelihood of any specific action is small. Preservation needs to be presented as a social good, one without which copyright holders would simply be unable to benefit from the property rights they seek to protect.  For UK institutions, the Jury recommends commentary by Naomi Korn on the status of orphan works and the impact of Brexit - that UK institutions are no longer able to make use of the EU		



Orphan Works Directive and the alternative Orphan Works Licensing Scheme is costly. A list of resources is available at <https://naomikorn.com/resources/>. For those in the UK, there is also the UK Copyright and Creative Economy Centre (CREATE) for resources on orphan works and copyright more broadly at <https://www.create.ac.uk/resources/>.

*Case Studies or Examples:*


- The National Disc of the BBC Domesday Project offers an example of loss outside of the legal deposit mandate where the copyright owner cannot be traced. See: <https://www.atsf.co.uk/dottext/domesday.html>


<b>PDF other than PDF/A</b> Documents presented in PDF (Portable Document Format) format (ISO 32000:1 and ISO 32000:2) and other data wrapped inside them, other than PDF/A but including all other variants and versions.		
Group: Formats	Trend in 2021:	
Added to List: 2017	<b>Trend towards reduced risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within five years, detailed assessment within three years.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> PDF 1.1, 1.2, 1.3, 1.4 (excluding PDF/A as a subset), 1.5, 1.6, 1.7 and 2.0. PDF/X and PDF/E		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Loss of context; loss of authenticity or integrity; external dependencies; poor storage; lack of understanding; significant diversity of data; poorly developed digitization specifications; lack of integrity checking; poorly developed migration or normalizations specifications; lack of virus control; poor storage or replication; lack of validation at the point of creation; encryption.		
<b>'Vulnerable' in the Presence of Good Practice</b> Well-managed data infrastructure; preservation planning; authenticity managed; use of persistent identifiers; reduction of dependencies; application of records management standards; recognition of preservation requirements beyond formats; strategic investment in digital preservation; preservation roadmap; participation in digital preservation community; format validation.		
<b>2021 Review</b> The 2019 Jury introduced this entry as a subset of a previous entry for 'PDF,' emphasizing the different threats faced by different types of PDF. PDF/A explicitly reduces dependencies and thus curtails preservation risks for certain types of content: PDFs of other types do not. PDF and PDF/A have sometimes been misunderstood as a generic solution to digital preservation requirements. In the Jury's eyes, it can only offer a preservation solution when embedded within a wider preservation infrastructure. The 2021 review agreed, noting a trend towards reduced risk as PDF continues to be a pretty stable format and there continue to be developments in tools and techniques (e.g., ability to convert PDF to PDFA to reduce dependencies). The <i>Endangered</i> classification remains given the need for support and embedding in a preservation infrastructure.		
<b>Additional Comments</b> There is a lot of material produced and kept in PDF. Some of it is authoritative, in other words, the only available copy, while some of it is not. However, if it is the only copy and it is lost, it can have an impact on a lot of people The challenge in evaluating the significance and impact of the loss of PDFs is that they're quite often a surrogate of something else, whether a digitized record or a Word document, etc. Whether or not that record is retained may be a factor. We should also be considering PDF Portfolios, which are an extension of PDF 1.7. Portfolios contain embedded files and can include		


text documents, spreadsheets, PowerPoints, emails, Computer Aided Design (CAD) drawings.

Assessing the risk of this complex format may need to be separate from other PDFs.


See also: Fanning, B (2017) *Preserving with PDF/A (Second Edition)*, DPC Technology Watch Report 17-01 online at <http://doi.org/10.7207/twr17-01>.

<b>Pre-Production TV and Movie Materials</b> Digital records of the creative and production process for film and television, such as initial designs, screenplay and script, on set still photography, rushes or out-takes that are not included in the final production and therefore not available to on-air broadcast archives or film libraries.		
Group: Sound and Vision	<b>Trend in 2021:</b>	
Added to List: 2017	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> TV and Movie production archives in digital form; outputs of script management software; drafts of a screenplay; continuity photography; costume design; set design; lighting and sound design.		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Lack of custodial responsibility; confusion over intellectual property rights; lack of appraisal; lack of recognition of preservation at executive level;		
<b>'Vulnerable' in the Presence of Good Practice</b> Preservation responsibility understood and acted upon; preservation infrastructure and planning for key items; access and use of collections to inform subsequent productions		
<b>2021 Review</b> This entry was first introduced in 2017 and noted as being 'of concern,' though the Jury did not have the capacity to assess the entry thoroughly. Additional expertise was recruited for the 2019 Jury, and the entry was added with the <i>Endangered</i> classification based on the materials not being collected in any coherent way, likely loss when not valued by production companies, costly space needed to hold them, and lack of expertise needed to catalogue and collect them. In other words, this is not primarily a technical problem and advocacy is needed urgently. The 2021 Jury added the trend towards greater risk in light of the rise of streaming services. The movie and film industry has been digitizing for a long time, but COVID has brought with it the rise of streaming services like Netflix, Disney+ etc., with productions that are tied to their organizations. These types of pre-production materials are great for promotion before and immediately after release; ongoing preservation after promotion finishes is unclear or uncertain.		
<b>Additional Comments</b> With the importance of advocacy in mind, it is important to raise awareness and educate directors and filmmakers about managing their digital archive so that it is still accessible if donated to a cultural institution in the future.		

<b>Proceedings in Court</b> Digital materials generated through legal proceedings in court.		
Group: Digital Legal Records	Trend in 2021:	
Added to List: 2017	<b>No Change</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to address losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> digital records of proceedings; digital records of rulings, and all manner of quasi-judicial proceedings and tribunals.		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Loss of context; loss of integrity; external dependencies; poor storage; lack of understanding; churn of staff; significant or diversity of data; poorly developed specifications; ill-informed records management; poorly developed transfer protocols; poorly developed migration or normalization; longstanding protocols or procedures that apply unsuitable paper processes to digital materials.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Well managed data infrastructure; preservation enabled at ingest; carefully managed authenticity; use of persistent identifiers; finding aids; well managed records management processes; recognition of preservation requirements at highest levels; strategic investment in digital preservation; preservation roadmap; participation in the digital preservation community.		
<b>2021 Review</b> This entry is a subset of a previous 2019 entry, ‘Proceedings and Evidence in Court,’ which was itself created as a subset of entry in 2017 for ‘Digital Legal Records and Evidence.’ The 2021 Jury split ‘Proceedings and Evidence in Court’ into two more discrete entries to highlight their distinct preservation challenges and risk profiles. This entry includes court proceedings and recognizes that courts have a responsibility to provide robust preservation that ensures the authenticity of these records.		
<b>Additional Comments</b> Standard Records Management processes within designated agencies should be able to take care of the preservation of materials like this, but given that it is likely to involve complex types of data, such agencies may not be equipped to deliver preservation effectively. It is surprising that courts are not more obvious in the digital preservation community, where solutions now exist.  Recordings of proceedings in court may include the AV recording of the court session, which may pose particular preservation risks associated with the video files.		

Published Research Data Appended to Journal Articles		
Closed research data sets produced and documented in accordance with good practice and simply appended to a journal article or transferred to a repository that does not have sufficient subject-matter expertise or funding commitment to ensure reliable or ongoing preservation for the long term.		
Group: Research Outputs	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards reduced risk	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Data sets added to papers in repositories that are designed primarily for papers; electronic journals offering data sets without obvious preservation capacity; institutional repositories servicing highly complex scientific data sets with insufficient subject-matter expertise.		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Unstable funding or revenues; poorly designed migration or normalization processes; poorly formed ingest and quality assurance procedures; rapid churn of staff; incoherent patterns of subject matter; lack of domain knowledge; no or very small numbers of users; weak or absent collecting policy; deposit to ensure minimal compliance with funder mandate; limited or dysfunctional data management planning.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Clear preservation planning; repository development roadmap; ability to transfer collections or share metadata with subject repositories or portals; strong user base; demonstrable re-use of data; clear collecting policy; data management planning early in the data lifecycle.		
<b>2021 Review</b> This 2019 entry was previously introduced in 2017 under 'Research Data,' though without explicit reference to the research data appended to journal articles. In 2019, the Jury split the 'Research Data' entry into a range of contexts for research outputs, including this addition. The entry draws attention to services that take upon themselves commitments to preserve research data, but which may not deliver those promises through lack of capability. The 2021 Jury agreed with the <i>Endangered</i> classification but commented on the improvements and initiatives towards preservation of research data outputs, with good practice documentation and replication in this space (e.g., collaborations with publishers and repositories, LOCKSS, CLOCKS, etc.). For these reasons, the 2021 trend is towards reduced risk.		
<b>Additional Comments</b> Research data is complex and has specific requirements for documentation which may only be known to subject matter experts. However well intended, it is risky for institutions to attempt to replicate that level of expertise across all the domains within the institution, and it can be hard for smaller publishers to make commitments to sustain data in the long term.		
Case studies or examples:		


- One example of publishers encouraging repository deposits is the FAIRsharing Collaboration with DataCite and Publishers. See: McQuilton, P., Sansone, S.A., Cousijn, H., Cannon, M., Chan, W.M., Carnevale, I., Cranston, I., Edmunds, S., Everitt, N. and Ganley, E., (2019) *FAIRsharing Collaboration with DataCite and Publishers: Data Repository Selection, Criteria That Matter*, online at <https://osf.io/m2bce/>.

<b>Recordings of Video Game Play Uploaded to Online Platforms</b>		
Recordings of game playing and e-sports that show how games are experienced and played, especially multi-user online games and tournaments.		
Group: Gaming	<b>Trend in 2021:</b>	Consensus Decision
Added to List: 2019	<b>Trend towards greater risk</b>	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Material uploaded to Amazon Twitch, game channels on YouTube and other playback services		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> Controversies around intellectual property rights; lack of offline backup; changing business model of providers; limited recognition of the cultural and historic value of game play; over-dependence on goodwill subsidy of an ad-hoc community; lack of preservation know-how at service providers; dependency on bespoke hardware or interfaces.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Offline backup; managed intellectual property rights; players and audiences invested in data		
<b>2021 Review</b> This entry was added in 2019 as a subset of an entry made in 2017 for ‘Gaming,’ which the Jury split into four more discrete entries. There are overlaps with the social media entries, except this category specifically draws attention to gaming and e-sports and therefore is a subset of both. By including as a separate entry, the 2019 Jury encouraged greater consideration of the cultural and historic value that such recordings are likely to acquire as well as the technical and economic challenges to preservation. The content is not particularly distinctive in technical terms, but there are aggravating circumstances, namely an almost complete reliance on commercial third parties (Google/YouTube and Amazon/Twitch) for the infrastructure around video capture and hosting. As the majority of this material is experienced and hosted on user-driven and ephemeral platforms such as YouTube and Twitch, it is less ‘collectable’ than the actual games and is unlikely to exist in private or public collections. Involves platforms that the digital preservation sector does not have much experience working with, e.g., Twitch. For these reasons, the 2012 Trend is towards greater risk.		
<b>Additional Comments</b> Copyright claims on video content by publishers such as Nintendo - while less prevalent now than a few years ago - also complicate things. The significance of loss here is high because recordings, including commentary, and onscreen interactions with other players, seem likely to be the best way of preserving the experience of playing certain games at certain times. We are familiar with the challenges of preserving video, but we need to think about how established approaches will work in the context of the aggravating circumstances outlined above. There is a degree of urgency		




associated with working out how (legally and technically) preserving the materials that they hold may be preserved.

Important for social context, and from a DP point of view videos should not be too hard, but if we are capturing the experience to inform digital preservation actions and intents, then do not these videos exist in places such as YouTube and wouldn't they be brought in as part of the 'documentary' evidence of DP actions taken on the game or sports that have come into the archive?

<b>Semi-Published Research Data</b> Data sets produced in the course of research and shared between researchers, such as by posting to a website or portal but without preservation capability or commitment. Typically the data remains in the hands of the researchers who have the job of maintaining it.		
Group: Research Outputs	Trend in 2021:	
Added to List: 2019	<b>Trend towards reduced risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Departmental webserver; project wikis; GitHub repositories		
<b>'Critically Endangered' in the Presence of Aggravating Conditions</b> Originating researcher no longer active or changed research focus; staff on temporary contracts; dependence on single student or staff member; weak or fluid institutional commitment to subject matter; weak institutional commitment to data sharing; complicated or contested intellectual property; encryption; limited or dysfunctional data management planning; web capture challenges that means unlikely to be picked up by automatic crawlers.		
<b>Vulnerable in the Presence of Good Practice</b> Data in preparation for transfer to specialist repository; robust data management planning; documented and managed professionally		
<b>2021 Review</b> This 2019 entry was previously introduced in 2017 under 'Research Data,' though without explicit reference to semi-published research data. In 2019, the Jury split the 'Research Data' entry into a range of contexts for research outputs, including this addition. The entry draws attention to represent 'self-help' data sharing which is to be encouraged as a means to facilitate open science but should not be confused with long-term preservation. The 2021 Jury agreed with the <i>Endangered</i> classification, noting problems with the volume of data being produced but not being kept in a meaningful way. Research data is complex and has specific requirements for documentation that may only be known to subject matter experts. However, data creators (e.g., researchers) are not necessarily well placed to sustain data in the long term. There have also been a few significant changes to the entry in the 2021 BitList. <ol style="list-style-type: none"> <li>1. Removal of 'informally' from the previous entry description ('shared informally between researchers') due to possible misperception or misunderstanding; 'informal' may imply researchers would perceive the data as low value and not want it captured. This may be the case, so it is important to consider and provide advice to researchers who think there is value in their data.</li> <li>2. Two previous entries (Geomagnetic Data and Software and Maritime Archaeological Archives) have been removed as separate entries and incorporated into this broader entry on semi-published research data to highlight the range of content and forms semi-</li> </ol>		

research data can take and highlight the need for specialized knowledge and specialist repositories for preparing and managing the data throughout the lifecycle
3. The 2021 trend towards reduced risk, based on improvements and initiatives towards the preservation of semi-published research data since the entry's addition in 2019.
<b>Additional Comments</b>
--

Virtual Reality Materials and Experiences		
Virtual reality (VR) refers to a set of technologies which build on existing 3D rendering technologies, with the aim of creating experiences which completely immerse a user in a virtual environment. The related term of Immersive Media (also known by the acronym XR) refers to a set of technologies used to create experiences, which either completely immerse a user in a virtual environment (Virtual Reality), augment the real world with virtual elements (Augmented Reality) or combine elements of the two (Mixed Reality). Key technologies include headsets, tracking systems, real-time 3D software and 360 video.		
Group: Media Art	New Entry	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance and Impact</b> The loss of tools, data or services within this group would impact on a large group of people and sectors.	<b>Effort to Preserve / Inevitability</b> It would require a major effort to prevent or reduce losses in this group, including the development of new preservation tools or techniques.
<b>Examples</b> Oculus Rift, VR tours, art installations		
<b>‘Critically Endangered’ in the Presence of Aggravating Conditions</b> lack of established frameworks and tools; technology is relatively poorly understood in the digital preservation domain; many of the technologies are proprietary; technology is seen as inherently fragile and therefore risky to collect and preserve; emulators do not currently support XR applications; expected to be difficult and costly to migrate, a process itself dependent on access to vulnerable source materials.		
<b>‘Vulnerable’ in the Presence of Good Practice</b> Effective replication; emulation; strong technical documentation; preservation pathway; good descriptive cataloguing; trusted repository.		
<b>2021 Review</b> This is a new 2021 entry submitted through the open nomination process. These VR technologies are finding use in many sectors, including archaeology, architecture, contemporary art, documentary film, gaming, forensics, science and engineering. While these technologies are not new per se, having experienced a first wave in the 1990s, they have experienced renewed interest recently as a result of a new generation of hardware. There are connections between this entry and others relating to both Media Art and Gaming, but it has been included as its own entry to emphasize the issues of preservation that pertain to the interconnected set of specific hardware and software components that access to XR experiences is contingent on. VR is challenging to document due to the individual nature of the experience, and components tend to become rapidly obsolete due to a fast rate of technological change as the industry pushes newer, higher fidelity hardware and software. This results in the potential to lose access to XR software applications, as old VR applications can no longer communicate with new XR hardware. The reliance on proprietary software and hardware components, as well as the lack of industry standards, poses a further risk.		
<b>Additional Comments</b> The current wave of materials made using XR technologies represent a unique point in time for the continued development of the technology and therefore represent a significant piece of		

computing history. Individual materials/experiences created using XR technologies present their own significance beyond this, which, noted elsewhere in this entry, can be represented in a wide range of sectors.

The impacts of the loss of access to virtual reality materials could be widely felt, given their wide-ranging uses across many sectors — most notably collections and archives containing materials accessed using these technologies. Simultaneously there is a risk of a loss of understanding of this technologies' development during the 2010-present period, which is likely to be of historical significance in and of itself.

*See also:*


- The Tate Preserving Immersive Media project, which is developing strategies for the preservation of artworks which utilize immersive media such as 360 video, real-time 3D, virtual, augmented and mixed reality, online at: <https://www.tate.org.uk/about-us/projects/preserving-immersive-media>
- Coates, C. (2021) Virtual Reality is a big trend in museums, but what are the best examples of museums using VR? MuseumXR, Museum Next, online at: <https://www.museumnext.com/article/how-museums-are-using-virtual-reality/>

## 10. Critically Endangered



Digital materials are listed *Critically Endangered* 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 *Endangered* materials in the presence of aggravating conditions and instances of *Practically Extinct* materials that have been identified but not fully recovered.

Community-generated Content in Arts and Heritage		
Digital materials produced and shared in and by ad-hoc community art and heritage projects, typically through digitization, where the creation of digital materials was a significant purpose of the initiative.		
Group: Community Archives	Trend in 2021:	Consensus Decision
Added to List: 2019	No Change	Previously: Critically Endangered
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Locally organized programmes associated with public remembrance and celebration such as World War One centennial commemorations; City of Culture; Olympic Games; World Cup		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Poor documentation; lack of replication; lack of continuity funding; lack of residual mechanism. dependence on a small number of volunteers, lack of preservation mandate; lack of preservation thinking at the outset; failure of digital legacy planning; conflation of backup with preservation; conflation of access and preservation; inaccessible to web archiving		
<b>‘Endangered’ in the Presence of Good Practice</b> Residual archive with residual funding able to receive and support collections; strict adherence to digitization guidelines; quality assurance; active user community; intellectual property managed to enable preservation.		
<b>2021 Review</b> This entry was added in 2019 as a subset of ‘Community Archives and Community-Generated Content’ which was split into two to provide greater specificity in recommendations. The 2020 Jury noted a trend towards greater risk based on how Community art and heritage groups, which often rely on volunteer effort, have been unable to meet for extended periods in 2020. Moreover, the local galleries, museums and arts centres on which they depend have closed, in some cases for good. Considering this dislocation, digital materials generated by community groups in arts and heritage are likely to be at a greater risk than in 2019. The 2021 Jury agrees with the <i>Critically Endangered</i> classification and 2020 trend towards greater risk. However, the risk has not increased to the extent where there is a notable 2021 trend. The Jury commented that risks have not increased as so much as the challenges have remained, specifically those of funding and sustainability. Community-generated materials are often produced and shared through funded projects and tell a similar story of loss through inaction, but the challenge is the same as before; sustainability with project-based funding. There should be greater Responsibility of funding bodies to ensure that digital preservation is built into any funded community preservation project plan and outcome. Digital content in this context is often seen as a by-product of engagement and is annoyingly left to disappear by organizations that do not have digital preservation experience or		


infrastructure. With good governance and sustainable digital repository support, this should not be an issue.

**Additional Comments**

Local archives address these collections on an ad hoc basis. This entry provides cultural insights into parts of society not addressed by collecting policies with a focus on the famous. Difficult to state whether national or international impact since it largely depends on the content and how widely it is used. I think this also depends on the identity group represented by the community group, such as events around Black History Month. Loss seems likely because of the precarity of the funding streams, or lack thereof, for these projects. Once digitization has been carried out, many projects do not know what to do with them or have the means to make them accessible. Often a lack of understanding of copyright is a barrier to sharing.

This may also be a call to education for grant funders about digital preservation of the long tail in funding these projects, such as what is happening now with government-funded research in universities and the universities realizing they have a responsibility to maintain the digital research data in a usable format.



<b>Consumer Social Media Free at the Point of Use</b> Social media services offered free at the point of use with a subscription model based on reselling user behavior and/or advertising.		
Group: Social Media	<b>Trend in 2021:</b>	
Added to List: 2017	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world	<b>Effort to Preserve</b> It would require a major effort to address losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Flickr, Vimeo, YouTube, Instagram, Periscope, DropBox, Facebook, Twitter, Pinterest, TikTok Yahoo Groups		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Lack of preservation capacity in provider; Lack of preservation commitment or incentive from provider; Lack of storage replication; proprietary products or formats; poor data protection; inaccessibility to web archiving; political or commercial interference; Lack of offline equivalent; super-abundance; poorly managed IPR; Lossy compression in upload scripts.		
<b>'Endangered' in the Presence of Good Practice</b> Offline backup and documentation of media assets; Migration plan; Early warning from vendors; Roadmap from vendors; Accessible to web harvest; Suitable export functionality; Licencing enables preservation; Preservation commitment from vendor; Preservation capability in vendor; Resilient to hacking; Selection criteria;		
<b>2021 Review</b> The 2019 BitList added this entry as a subset of a previous 2017 entry, emphasizing the different threats faced by online services that are 'paid for versus 'free at the point of use.' Both depend on the business model of the vendor and the terms which they impose. There are overlaps with the 2021 entry 'Cloud-based Services and Communications Platforms', but this entry remains separate to highlight that for this group of social media services as digital materials, the business model and sustainability can only be guessed, and contracts tend to be asymmetrical in favour of the supplier. Moreover, because these services have a low barrier to entry, they may be favoured by agencies or individuals least able to respond to closure or loss. Preserving this stuff en mass is still incredibly difficult, but many of these platforms allow the downloading of their own personal content/archives. However, these lose all the context of social media and therefore, whilst they do preserve the data, they do not preserve the essence of the material. Platforms like Twitter have opened their API further in recent years, but others like Yahoo have closed, and Facebook continues to be almost hostile towards archiving and preservation attempts. Over the last year, concerns have arisen with a trend towards harmful and malicious hate speech and misinformation, deliberate deletion. For these reasons, there is a 2021 trend towards greater risk but no change to the <i>Critically Endangered</i> classification.		


### **Additional Comments**

Social media capture via web harvesting has become increasingly difficult. The platforms continue to put up barriers to automated capture that prevent the preservation of even so-called public content. For example, campaign websites or other election-related content that is only published on Facebook or on Twitter because these services are 'free.' This content is of particular concern as it appears on no other website. Web archivists are constantly shifting strategies and approaches and trying out new (but limited) tools to best capture this content. If we cannot successfully preserve these platforms, we are missing out on documenting organizations, campaigns and elections around the globe. Much of this data exists as data sets based on aggregated use rather than individual files.


Often these are external proprietary platforms bound by intellectual property law and potentially privacy law which will impede the imminence of action. What recourse do archives or digital repositories have to deal with this and capture the materials?

#### *Case Studies or Examples:*


- An example of a tool available to help libraries and archives with capture is *Archive Social*, online at <https://archivesocial.com/>.
- A range of use cases are presented in Thomson, S. (2016). *Preserving Social Media*, DPC Technology Watch Report (16-02), online at <http://doi.org/10.7207/twr16-02>.

<b>Correspondence and Records of Research</b> Correspondence and other records which describe the configuration and delivery of research but which are ancillary to the core research outputs, including reviews, drafts and correspondence between researchers.		
Group: Research Outputs	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment in one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Email boxes of senior academics; social media posts; personal spaces on institutional networks		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Originating researcher no longer active or changed research focus; staff on temporary contracts; dependence on single student or staff member; weak or fluid institutional commitment to subject matter; weak institutional commitment to data sharing; complicated or contested intellectual property; encryption		
<b>'Endangered' in the Presence of Good Practice</b> Recognition of value of correspondence; integration with CRIS; routine use of EDRMS; documented and managed professionally; separation of personal and corporate identities		
<b>2021 Review</b> This entry was introduced in 2017 under 'Research Data,' though without specific reference to the correspondence and records of research. In 2019, the Jury split this entry into a range of contexts for research outputs. There was a 2020 trend towards greater risk based on education and research institutions facing budget uncertainties, and a number of institutions have introduced early severance schemes or put staff on short term contracts at greater risk of redundancy and, while this puts other types of research output at risk, the personal nature of correspondence means that the risks are intensified, and so this item faces greater risks than identified in 2019. However, the 2021 Jury agreed there had been no significant change in the trend over the last year, but this may change in the future with more examples, such as those pertaining to Covid.		
<b>Additional Comments</b> In an ideal circumstance, correspondence should be stored in EDRMS systems separately from research data and subject to different retention schedules, i.e., 10-20 years. There may be challenges connecting the EDRMs holdings to the research data and vice versa.  Advocacy and research re the scale of the problem may be required to encourage academics to use EDRMs, for example, correspondence and integration with CRIS. Simplified tools and workflows to move data from CRIS to Repository to Preservation systems		

There should also be an encouragement to researchers to keep only what is needed and only for so long as the retention period requires. More often than not, records of correspondence will not require long-term preservation.

<b>Data Posted to Defunct or Little-used Social Media Platforms</b> Older or less widely used social media platforms to which content has been uploaded but for which no guarantees have been made about the long term		
Group: Social Media	<b>Trend in 2021:</b>	
Added to List: 2017	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment in one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> BeBo, MySpace, Google Buzz and others		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Closure of platform; lack of offline equivalent; lack of export functionality; no preservation undertaking from service provider; unstable business plan from service provider.		
<b>'Endangered' in the Presence of Good Practice</b> Offline Replication; clear notice periods and alerts; committed ongoing maintenance of service		
<b>2021 Review</b> The 2019 judges added to this entry introduced in 2017, emphasizing the different threats faced by social media users who uploaded content to defunct or little-used social media platforms. Because these services are older, the need to act is more urgent than for others. Often, the significance is only brought to attention once they are lost, which in some instances may pose the question of the to preserve them. When we look at the digital preservation landscape and where we need to apply effort as well as resource defunct early social media spaces are not high on the list, but when considering how contemporary social media channels could become defunct, it becomes a different conversation because of how intrinsically tied they are to political discourse and influencing political opinion. For this reason, there is a 2021 trend towards greater risk due to the existing risks of defunct or little-used platforms with recognition of the need to develop tools or techniques for applying to others that may follow the same path.		
<b>Additional Comments</b> It is to be hoped that some of these have been archived via traditional web archiving, and so the remnants of these sites can be found in bits and pieces in various web archives, but it may be too late to save some of the content that is likely already be lost. If some of this is still available, there may be hope in trying to preserve, but it may be difficult if the platforms are not willing to share data or work with preservationists. ArchiveTeam has stepped in here too. There is undoubtedly a story here that could be used as a call for arms to raise awareness about the preservation of current social media platforms too		
<b>Case Studies or Examples:</b> <ul style="list-style-type: none"> <li>The shutting down of Yahoo Answers is an example of the loss of content from community board style web platforms. See: Tsavkko Garcia, R. (2021) <i>Deleting Yahoo</i></li> </ul>		

*Answers is a disastrous idea. For history's sake, we need to preserve our digital record.*  
Business Insider, <https://www.businessinsider.com/deleting-yahoo-answers-disastrous-idea-preserve-our-digital-record-2021-4?r=US&IR=T>.


Digital Archives from Public Enquiries and Commissions		
Data from public enquiries and reconciliation commissions which can be traumatic, politically uncomfortable and contested, typically comes in many different forms and formats. Data protection issues and cultural sensitivities only amplify the challenge to preservations.		
Group: Digital Legal Records	Trend in 2021:	Consensus Decision
Added to List: 2017	No Change	Previously: Critically Endangered
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> The Tunisian Truth and Dignity Commission to investigate human rights violations committed prior to 2012; enquiries into historical child abuse; Bloody Sunday Enquiry (Saville Inquiry); East Timor Tribunal.		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Risk of falsification; fragile or obsolete media; dependence on proprietary formats or products; lack or loss of documentation; inaccessible to web harvesting technologies; lack of version control; lack of integrity checks or integrity records; poor chain of custody; inability to identify an archival authority		
<b>‘Endangered’ in the Presence of Good Practice</b> Strong sense of archival responsibility; carefully constructed rules around information privacy that retain robust and appropriate preservation capabilities; clear legislation on retention and permanency, an appraisal of perceived value with resources to undertake preservation actions		
<b>2021 Review</b> In 2019, this entry became a subset of an entry introduced in 2017 for ‘Digital Legal Records and Evidence,’ which was split into four more discrete entries. This category includes evidence from public enquiries and commissions that have been presented in court. It recognizes that courts are not limited in the types of evidence that they can admit but that they have a responsibility to provide robust preservation that ensures the authenticity of their records and evidence. The 2021 Jury noted that there is considerable evidence of good practice emerging from some of the examples where clear archival responsibility has been the key to progress. National or state recordkeeping regimes and legislation are pretty clear on the retention or permanency of these types of records. A major issue, however, is embargoes. When an embargo is lifted, will the file format or database continue to work, or will it longer work, making the data useless?		
<b>Additional Comments</b> Case files and correspondence are one thing. Retention of these should be clear but may differ widely between jurisdictions and levels of government. If retention is not long-term or permanent, the risk of loss may not be so critical. Retention of 'unused' or 'potential' evidence is likely a different matter altogether. It may not even be considered a record, and certainly is not a record of the court. Should it be returned to the suspect or accused? Are their rights being		

considered here - not just in terms of preservation, but also simply disposition? There are legal and ethical issues around this that need to be fleshed out in conjunction with assessing its preservation risk.


*Case Studies or Examples:*


- The Hillsborough disaster and inquiry as a case example of issues arising when inquiry records get split up. A copy of *The Report of the Hillsborough Independent Panel* is available online at [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/229038/0581.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/229038/0581.pdf); see also Sarsfield, D., (2021) *Hillsborough: The Archive as a Vehicle for 'Truth Recovery.'* British Online Archives. <https://microform.digital/boa/posts/category/articles/423/hillsborough-the-archive-as-a-vehicle-for-truth-recovery>
- The Tunisian Truth and Dignity Commission, (2019) *Tunisia: Truth Commission Outlines Decades of Abuse*. Human Rights Watch, <https://www.hrw.org/news/2019/04/05/tunisia-truth-commission-outlines-decades-abuse>
- Bloody Sunday Enquiry (The Saville Inquiry), *The Saville Inquiry: Q&A*. BBC News, <https://www.bbc.com/news/10147362>
- East Timor Tribunal, Special Panels for Serious Crimes (East Timor), copy of the *Truth and Dignity Commission's (IVD) final report* available in English at [http://www.ivd.tn/rapport/doc/TDC\\_executive\\_summary\\_report.pdf](http://www.ivd.tn/rapport/doc/TDC_executive_summary_report.pdf)




<b>Digital Archives of Community Groups</b> Digital materials including ephemera, correspondence and campaign materials created as a by-product of small scale or ad-hoc community action groups		
Group: Community Archives	<b>Trend in 2021:</b>	
Added to List: 2019	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Archives of smaller and ad-hoc political and campaigning organizations; environmental protests; sports clubs; smaller religious groups; amateur music or drama; fan groups		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Poor documentation; lack of replication; lack of continuity funding; lack of residual mechanism; dependence on small number of volunteers, lack of preservation mandate; lack of preservation thinking at the outset; conflation of backup with preservation; conflation of access and preservation; inaccessible to web archiving; dependence on social media providers; distrust of 'official' agencies.		
<b>'Endangered' in the Presence of Good Practice</b> Residual archive with residual funding able to receive and support collections; active user community; intellectual property managed to enable preservation.		
<b>2021 Review</b> The jury created this entry in 2019 as a subset of 'Community Archives and Community-Generated Content' which was split into two to provide greater specificity in recommendations for approaching the preservation of created as a by-product of small scale or ad-hoc community action groups (versus digital materials generated for significant purpose of a community initiative). There was a 2020 trend towards greater risk based on community groups such as sports clubs, religious communities, arts and political groups, often relying on volunteer effort, being unable to meet for extended periods in 2020. Moreover, the local community centres, clubs or places of worship on which they depend have closed, in some cases for good. This trend continues for 2021; the Jury commented that much of the content in community archives has easily preservable content, just the resources are not directed towards them, basic digital preservation practices are not well embedded amongst the general population, and selective approaches are needed to get a handle on the situation and to find the resources to do the work.		
<b>Additional Comments</b> Typically born digital material is more at risk - community groups may not know about the risk of loss. Many are unaware of digital preservation terminology. It is the ad-hoc nature of these groups and projects which is of great concern.		


Significant need to raise awareness and provide a 'home' but also to do so with sufficient sensitivity so as to ensure community groups remain in control of their own material.

<b>Digital Archives of Music Production</b> Digital materials created by musicians and fans as a by-product of performance or recording, not otherwise published or shared		
Group: Sound and Vision	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent losses in this group, such as the development of new preservation tools or techniques.
<b>Examples</b> Pre-production notes; demo recordings; photography; correspondence.		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> fragile or obsolete media for offline content; service provider preservation capability for online content; dependence on proprietary formats or products; lack or loss of documentation; uncertainty over intellectual property rights; lack of version control; lack of policy or mandate		
<b>'Endangered' in the Presence of Good Practice</b> Replication; clarity of intellectual property rights; preservation agency involved and capable of looking after content		
<b>2021 Review</b> In 2019, this entry was created as a subset of a previous 2017 entry, 'Digital Music Production and Sharing,' which was split into to draw attention to the different challenges faced by the different forms. Although it has overlaps with other entries, including 'Pre-production TV and Movie materials,' it is a separate entry to emphasize the inherent and value of the archival materials relating to the recording process over and above the recordings themselves.		
<b>Additional Comments</b> For imminence of action, it will all depend on the format of the records. Correspondence or photographs may be left for longer, but recordings will need closer attention, especially if it is a bespoke recording format. This may be less of digital preservation challenge and more of an archive or collecting challenge. This type of material in the past, like most 'unpublished' archives, has survived through luck and is largely out of a GLAM or institution's control relying on individuals to assess and evaluate if what they have is of significance. When these types of things come to an institution, based on significance on a case-by-case basis, are these digital objects then considered 'worth' the effort to a) bring into the collection and b) care and preserve them? So yes, while <i>Critically Endangered</i> , these types of collections are enormous and quite often not things one would want to keep for the long term; however, sometimes there is the odd gem. There is the recognized inevitable loss of existing data, but reducing this loss would require major effort to fix in terms of identifying organizations who are preserving this content, and it is not clear that this is being done already.		

<b>Digital Evidence and Records of Investigation Prior to Court</b> Digital materials assessed by police and other authorities in the course of investigation and retained as evidence of due process such as case files and correspondence, including materials not submitted to court		
Group: Digital Legal Records	<b>Trend in 2021:</b>	
Added to List: 2019	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent losses in this group, such as the development of new preservation tools or techniques.
<b>Examples</b> CCTV; Email; 3d scanning; social media interactions; police records; court records; text messages.		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Poor chain of custody; fragile or obsolete media; dependence on proprietary formats or products; lack or loss of documentation; inaccessible to web harvesting technologies; lack of version control; lack of integrity checks or integrity records; poor chain of custody.		
<b>'Endangered' in the Presence of Good Practice</b> Meticulous transfer and disclosure processes		
<b>2021 Review</b> This entry was added in 2019 as an entry made in 2017 for 'Digital Legal Records and Evidence,' which the Jury split into four more discrete entries. This category includes evidence prior to court that may form part of an investigation or gathering of evidence but which are not formally submitted as evidence. It recognizes that police and other investigating authorities are not limited in the types of evidence that they need to administer, but that this creates an almost unbounded limit of preservation requirements to ensure authenticity and admissibility. A 2021 risk was identified based on examples bringing to question whether legal bodies have the skills and capabilities to preserve these materials should they need them if a case is reopened etc.		
<b>Additional Comments</b> Case files and correspondence are one thing: retention of these should be clear but may differ widely between jurisdictions and levels of government. If retention is not long-term or permanent, the risk of loss may not be so critical. Retention of 'unused' or 'potential' evidence is likely a different matter altogether. Is it even a record? Certainly, it is not a record of the court. Should it be returned to the suspect or accused? Are their rights being considered here - not just in terms of preservation, but also simply disposition? There may be legal and ethical issues around this that need to be fleshed out in conjunction with assessing its preservation risk.		


<b>Evidence in Court</b> Digital materials presented in court as evidence or documents such as rulings and proceedings generated through legal proceedings		
Group: Digital Legal Records	<b>Trend in 2021:</b>	
Added to List: 2017	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to address losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Evidence submitted to courts of all kinds, including text messages, photography, CCTV, email, 3d and 2d scanning, scientific reports and analyses, documents and websites;		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Loss of context; loss of integrity; external dependencies; poor storage; lack of understanding; churn of staff; significant or diversity of data; poorly developed specifications; ill-informed records management; poorly developed transfer protocols; poorly developed migration or normalization; longstanding protocols or procedures that apply unsuitable paper processes to digital materials.		
<b>‘Endangered’ in the Presence of Good Practice</b> Well managed data infrastructure; preservation enabled at ingest; carefully managed authenticity; use of persistent identifiers; finding aids; well managed records management processes; recognition of preservation requirements at highest levels; strategic investment in digital preservation; preservation roadmap; participation in digital preservation community.		
<b>2021 Review</b> This entry is a subset of an entry made in 2019 titled ‘Proceedings and Evidence in Court,’ which was itself created as a subset of entry in 2017 for ‘Digital Legal Records and Evidence,’ The 2021 Jury split ‘Proceedings and Evidence in Court’ into two more discrete entries to highlight their distinct preservation challenges and risk profiles. This entry includes evidence that has been presented as evidence in court. It was given a <i>Critically Endangered</i> classification to highlight its higher risk profile and additionally emphasize that courts are not limited in the types of evidence that they can admit but that they have a responsibility to provide robust preservation that ensures the authenticity of their records and evidence.		
<b>Additional Comments</b> Standard Records Management processes within designated agencies should be able to take care of the preservation of materials like this, but given that evidence is likely to involve complex types of data, such agencies may not be equipped to deliver preservation effectively. It is surprising that courts are not more obvious in the digital preservation community, where solutions now exist.  More concrete examples would be welcome. It is the evidentiary value of submissions to court that may be lost, and therefore veracity of the decision could be questioned. Evidence submitted		

in digital form is of greater risk (e.g., a video file submitted on a CD in the 90s) than records of the proceedings themselves (e.g., transcripts).

<b>Family or Personal Records</b> Digital content and communications generated for personal consumption in a domestic setting and which may be of limited general interest but highly valuable to family members and genealogy.		
Group: Personal Archives	Trend in 2021:	
Added to List: 2017	<b>No Change</b>	Unanimous Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of data, tools or services within this group would have a localized impact.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Childhood photographs and videos; School or graduation photos; wedding photos and movies; electronic correspondence (email, messenger, WhatsApp)		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Storage on portable media or poor storage; dependence on devices or processes; dependence on obsolete or proprietary formats; storage media out of warranty; single copies; inappropriate dependence on service provider; inappropriate encryption or password protection; lack of awareness or planning; loss or lack of documentation; over-abundance; inability to act in a timely manner; confusion over intellectual property		
<b>'Endangered' in the Presence of Good Practice</b> Replication; action in a timely manner; open formats; selection and appraisal; archival agency		
<b>2021 Review</b> This was introduced to the BitList in 2017. Although research and advice on the preservation of personal records have been available for some time, outreach and training have not reached the audience, and there has been no material improvement in the risks faced by this category since 2017. It is reasonable to assume that the number of digital objects in this category has increased; thus, the consequences of loss have expanded but the 2021 Jury determined there has been no significant trend towards greater risk; content is being lost all the time despite digital materials that can easily be preserved with tools not widely available outside of institutions. Therefore, this is a public awareness campaign issue and more tools need to be made easily available for people to be able to better preserve their own digital content.		
<b>Additional Comments</b> There is a strong overlap with community archives, except noting that responsibility is even more localized. There is room breaking the entry down further into a series of components to represent the complexity more effectively and present a more nuanced action plan.  This matter needs awareness-raising. Education is needed, such as digital preservation as a survival skill for teenagers. Also, simple and cheap tools or pathways to preservation are needed.  Education to the public is critical for advocacy - these are the societal records of the future! Though having said that, what has survived in hardcopy has largely been through luck, and the		


same thing I think will be the same for digital. The same issues exist with glass plate negatives, photographs and certain emulsions and even printed digital photographs, brittle paper, fading ink etc.




<b>Grey Literature</b> Semi-published research outputs such as blogs, dissertations, informal conference papers or commissioned reports which are not formally published but which can contain original and insightful contributions within scholarly communications		
Group: Research Outputs	Trend in 2021:	
Added to List: 2019	<b>Trend towards reduced risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Blogs, technical reports, conference papers, dissertations, commercial research		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Originating researcher no longer active or changed research focus; staff on temporary contracts; dependence on single student or staff member; weak or fluid institutional commitment to subject matter; weak institutional commitment to data sharing; complicated or contested intellectual property; encryption; Lack of recognition; non-disclosure agreements;		
<b>'Endangered' in the Presence of Good Practice</b> Use of persistent identifiers; embedded within repository infrastructure; quality assurance		
<b>2021 Review</b> This entry was introduced in 2017 under 'Research Data,' though without explicit reference to grey literature. In 2019, the Jury split this entry into a range of contexts for research outputs. This entry represents activities which build towards formal publications and research outputs but which do not typically accumulate in institutional repositories. The 2021 Jury agreed; however, there was a significant difference between the 2020 trend and the 2021 trend. The 2020 Jury noted a trend towards greater risk because higher education and research institutions faced budget uncertainties, and a number of institutions have introduced early severance schemes or put staff on short term contracts at greater risk of redundancy; While this puts other types of research output at risk, the ad hoc nature of grey literature means that this entry is at greater risk. Members of the 2021 Jury argued the content of grey literature is not entirely unique if it eventually makes its way into published outputs and noted improvements and initiatives towards preservation of semi-published research data and outputs over the last year, resulting in the consensus of a 2021 trend towards reduced risk.		
<b>Additional Comments</b> This is why it is called grey literature - it is the same in the analogue world - publications come and disappear. Loss of material like this would be common in the analogue world, but in the digital age, we have the capacity and perhaps something of a responsibility to ensure that it is captured: more of an opportunity lost to extend the available research resource. The ADS's Grey Literature Library		

demonstrates what could be done if information architectures are deployed to mirror and extend professional practice.


Workflows and policies re tagging, collecting and EDRMS may help protect such data into the future. Past materials are almost certainly partially lost

Legacy Research Web Collections		
Since 1994, people have set up collections of digital content on the web with software now outdated. Those collections are valuable but lose funding and care as institutions re-configure their tasks and individuals retreat from their task due to retirement or (as volunteers) to old age.		
Group: Web	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards greater risk	Previously: Critically Endangered
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Academic and institutional website from the first decade of the web containing details of research projects and interests as well as research data.		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Inaccessible to web archive; bespoke code; insufficient documentation; uncertain intellectual property right.		
<b>‘Endangered’ in the Presence of Good Practice</b> Secured by web archive; documentation and rights information published alongside		
<b>2021 Review</b> This entry was added in 2019. There are overlaps with the entry with the ‘Semi-Published Research Data’ entry, but it is a separate entry to distinguish between ‘current’ and ‘legacy’ collections with different risk profiles: in this case, the fact that materials of legacy web collections are no longer actively maintained increased the classification to <i>Critically Endangered</i> . Loss has already occurred, and future loss can be prevented through approaches such as web archiving and code preservation, but the risk has become greater due to security issues posed by hosting legacy technology software and services which have prompted disposal of content imminently without adequate review or selection.		
<b>Additional Comments</b> The Internet Archive and other national web archiving bodies have copies of a lot of websites that would fit into this category but by no means all. There’s also a distinction between the software or code used to deliver the user experience and the data. Such code is secondary to the content.  This issue can be intensified by the legacy IT Infrastructure in cases where much of the content is hosted there, as security concerns may lead to disposal of content imminently IT. In these scenarios, their imminence of action becomes more urgent given the security issues posed by hosting legacy technology/software/etc.		


Massively Multiplayer Online Gaming Platforms and Experiences		
Massively Multiplayer Online (MMO) Gaming is an evolving, transient but significant culture. Gameplay is referenced here particularly as means of participation, along with social media and in-game interaction between players. Video streaming of game content, the means of viewing but not participating in game play, has a separate entry.		
Group: Gaming	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards greater risk	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples:</b> There are numerous examples for this entry, so the example of Fortnite, 2017 free-to-play online game will serve to illustrate. As Fortnite is freely accessible to play, the game is at the forefront of internet culture, with over 250 million registered users as of March 2019. In addition, Fortnite's parent company, Epic Games, provides access to Unreal Engine, a software platform consisting of integrated tools for game developers, with instructional guides, game simulations and learning tools. Unreal Engine was first released in 1998 and is currently in its 4th release. At risk also is this software, its associated tools and code, which is key digital evidence of MMOG evolution, learning, creativity and expression across multiple platforms (PC, console, mobile).		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Controversies around IPR; lack of offline backup; changing business model of providers; limited recognition of the value of game play; over-dependence on goodwill of ad-hoc community; lack of preservation know-how at service providers; dependency on bespoke hardware or interfaces.		
<b>'Endangered' in the Presence of Good Practice</b> Well documented code; IPR supportive of preservation; large and committed user community		
<b>2021 Review</b> This entry was added in 2019 as a subset of a 2017 entry for 'Gaming' which was split into four discrete entries. By creating this entry for Game Play, the Jury encouraged greater consideration of the technical complexities which arise from preservation of software environments as well as the cultural and historic value which these games are likely to acquire. The 2020 Jury noted a trend towards greater risk because online gaming has been a mainstay of social interaction during the pandemic. The risks of loss have not strictly increased, but the significance has expanded, so the consequences of loss are expanded. The 2021 trend also reflects this. The 2021 Jury agreed with elevating classification to <i>Critically Endangered</i> as 1) many multiplayer online games may already be lost as servers have been taken offline 2) those that remain tend to change rapidly, resulting in loss of access to early iterations of the experience (e.g., the frequent changes made to the virtual worlds of World of Warcraft and Fortnite). I wonder if it would be worth changing this to dropping the 'Massively' to ensure a broader range of multi-player gaming experiences are represented (e.g., smaller-scale co-op games). These are likely to become increasingly significant as games are more frequently integrating an optional online component (e.g., the Dark Souls series).		
Additional Comments		

The difficulty of saving game play nicely encapsulates why video recordings of (online) gameplay are important: we are never going to be able to recreate the experience of playing something like World of Warcraft or Fortnite at their peak (or at their inception, or at any other point in time). They will never have the same configuration of subscribers, to say nothing of the innumerable changes made to the software over the years, which have significantly altered how the game works and looks. Loss is inevitable, and it has already happened. The social and cultural aspects of play are incredibly important, and on-screen recording is the most robust way to capture that.


There may be the need for current clarification when talking about the connection between all the media and formats and the experience or just the experience?


<b>Media Art by Deceased Artists or Defunct Workshops</b>		
Media art where the artists or creative technicians are either deceased or not able to provide guidance on authenticity and installation		
Group: Media Art	<b>Trend in 2021:</b>	Consensus Decision
Added to List: 2019	<b>Trend towards greater risk</b>	Previously: Critically Endangered
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Works produced by media artists now deceased, such as: Jeremy Blake, Beatriz Da Costa, Heiko Daxl or Stanislaus Ostoja-Kotkowski.		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Lack of documentation to enable maintenance; lack of clarity with respect to intellectual property; complex interdependencies on specific hardware, software or operating systems; lack of capacity in the gallery or workshop; lack of strategic investment; complex external dependencies; loss of institutional memory resulting from staff churn; poor working relationship between the gallery and artist/workshop; lack of conservation assessment.		
<b>‘Endangered’ in the Presence of Good Practice</b> Strong documentation; clarity of preservation path and ensuing responsibilities; proven preservation plan; capacity of workshop to support re-installation; capacity of gallery to conserve; capacity of gallery to re-install; retention of institutional memory including archives of correspondence between gallery and artist/workshop; strong and continuing working relationship between the gallery and artist/workshop; regular conservation assessment.		
<b>2021 Review</b> This entry was added in 2019 as a subset of the 2017 ‘Media Art,’ which was first introduced with particular reference to historical media art but split by the 2019 Jury to ensure greater specificity in its recommendation. This entry represents works held in galleries where the artist is deceased or the workshop has closed, and there is limited prospect to obtain new documentation. The 2020 Jury found a trend towards greater risk based on how galleries, which often rely on visitors for income, have been closed for extended periods and circumstances of economic dislocation. The 2021 jury agreed on a continued trend towards greater risk based on the increasing risk of this loss happening with more time sensitivity for early media artworks.		
<b>Additional Comments</b> This entry includes a point in the lifecycle of all media art, so good practice recommendations are likely to become more important over time. Preservation issues may not become visible until the piece is brought out of storage for loan or exhibition, underscoring the value of continuous or periodic conservation assessment. The range of data/formats/hardware/software etc. can be new and varied, providing organizations with an ongoing technical challenge that they are not initially equipped to deal with. Some loss seems inevitable.		


We need to learn what documentation and metadata we need. We have an example where we have extracted art files, but we do not know what the artist's intent was, and we cannot ask as he has died.


<b>Media Inside Paper Files</b> Media inside paper files occurred in records since the 1980s and will continue to do so for many years.		
Group: Portable Media	<b>Trend in 2021:</b>	
Added to List: 2019	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Digital media mixed with paper files in records offices and filing cabinets of almost every kind of enterprise.		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Unsustainable effort to assess; exotic or obsolete media; poor storage; lack of descriptive labelling;		
<b>'Endangered' in the Presence of Good Practice</b> Carefully labelled; managed programme of assessment and retrieval; robust media used		
<b>2021 Review</b> This entry was added in 2019 to report the significant amounts of digital media being transferred to archives folded into traditional files. The 2019 Jury noted that it is relatively simple to preserve this material once identified using standard tools, but it can be an 'unknown unknown,' and that assessment can seem overwhelming and, therefore it may overlap with other portable media risks but has a higher risk classification. The 2021 Jury agreed on a trend towards greater due to the increased time sensitivity and need for conducting collection audits as soon as possible determine what you have to then work out a plan about opening carriers, assessing files and extracting them if significant.		
<b>Additional Comments</b> Highly dependent on who is looking after the portable formats. There are good examples, for example in libraries, where disks are stored at the back of books or front of magazines and can be processed at the point of acquisition. In archives, however, dealing with bit-level preservation of external media (often on legacy formats) is largely an unquantified problem, and so resource commitments will not be in place. So there is a method and tools but simply no time committed and no proper assessment either. In other agencies, the issue will not have even been considered, and for them, it will be much harder over time with some inevitable loss.  I would say media in paper files dates to earlier than the 1990s, floppy disks really back to the 1970s		




Non-current Hard Disk Technologies		
Materials saved to storage devices with a variety of underlying magnetic or solid-state technologies that are hardwired into a computer that is no longer under warranty or supported: typically, hard disks more than five years old.		
Group: Integrated Storage	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards greater risk	Previously: Critically Endangered
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Disks installed into computers or servers that are more than five years old, or out of warranty		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Lack of replication; poor storage; non-standard connections or controllers; aggressive compression; encryption		
<b>‘Endangered’ in the Presence of Good Practice</b> Maintenance schedule; renewable extendable warranty; best practice storage and operation; replication		
<b>2021 Review</b> This entry was added in 2019 to ensure that the range of media storage is properly assessed and presented. The lifecycles of most consumer hard disk technology is stable in comparison to portable devices because they are integrated into systems and therefore inherit the lifecycle and replacement of the entire system. This is less true at scale; however, where disks are used in storage arrays, and refreshment is more loosely tied to the server architecture. Storage at scale also means the percentage likelihood of finding a disk failure increases, and this likelihood of failure led to the 2021 Jury’s noted trend towards greater risk.		
<b>Additional Comments</b> The greater density of newer disks, as well as encryption and compression, mean they can be more fragile than older disks with less density, and less sophisticated read/write technologies. The age of a disk is not the best or only indicator of its reliability.		


<b>Non-current Portable Magnetic Media</b> Materials saved to floppy disks, tape, portable hard disks or other magnetic storage devices where the media is out of warranty and reader devices may no longer be supported or integrated easily into hardware infrastructure: typically, more than five years old.		
Group: Portable Media	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Unanimous Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment in one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Floppy disks; tape; certain kinds of portable hard disks, zip drives.		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Poor storage; inability to access readers; no replication; encryption; aggressive compression		
<b>'Endangered' in the Presence of Good Practice</b> Active management; dependable access to readers; strong documentation; documentation independent from the media		
<b>2021 Review</b> The judges have introduced this entry to ensure that the range of media storage is properly assessed and presented. Portable magnetic media was ubiquitous but is fragile not just to physical wear and tear but also to magnetic interference and bit-rot. The substrates of the disks can prove unstable, and in some cases, proprietary reader technology means that the disk becomes obsolete before it degrades. Storage at scale also means the percentage likelihood of failure increases.		
<b>Additional Comments</b> We know what to do with this type of material, it is the scale that makes it a problem There is really no excuse for using floppy disks for storage these days. Tape is a different proposition since it allows high-density back up offline and nearline. But there are challenges with the backward compatibility of popular and even relatively recent LTO versions.		

<b>Non-current Portable Optical Media</b> Materials saved to DVDs, CDs or other optical storage devices where the media is out of warranty and reader devices may no longer be supported or integrated easily into hardware infrastructure: typically, more than five years old.		v 
Group: Portable Media	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Unanimous Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment in one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> CDs, laserdisc technologies, DVD, HDVD		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Poor storage; inability to access readers; no replication; encryption		
<b>'Endangered' in the Presence of Good Practice</b> Active management; dependable access to readers; strong documentation; documentation independent from the media		
<b>2021 Review</b> This entry was added in 2019 to ensure that the range of media storage is properly assessed and presented. Optical media is in some senses a preferred option as it is typically more stable than magnetic or solid state media: but these characteristics have been over-sold. The substrates of the disks can prove unstable, and more importantly, the reader technology tends to be proprietary and can become obsolete long before the disks degrade. Storage at scale also means the percentage likelihood of failure increases.		
<b>Additional Comments</b> We know what to do; it is the scale of the problem. It is a big scale problem - and in many library catalogues the information about these carriers and their playing requirements do not exist in metadata. Big scale surveys of collections to identify these carriers might be a barrier to starting to tackle this one.  Highly dependent on who is looking after the portable media but made more difficult over time. There is a lack of granularity in the definition of media types, so generic advice like 'make sure you're moving your data to new forms of storage every 5 -10 years'.  Early generations of this media are not as robust as the current generation and can deteriorate significantly if not stored appropriately.		

<b>Non-current Portable Solid-State Media</b> Materials saved to flash or other solid-state storage devices where the media is out of warranty and reader devices may no longer be supported or integrated easily into hardware infrastructure: typically, more than five years old.		
Group: Portable Media	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Unanimous Decision
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment in one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> USB sticks and pen drives; Flash storage in cameras and phones; certain types of portable hard disk		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Poor physical storage; inability to access readers; no replication; encryption		
<b>'Endangered' in the Presence of Good Practice</b> Active management; dependable access to readers; strong documentation; documentation independent from the media		
<b>2021 Review</b> This entry was added in 2019 to ensure that the range of media storage is properly assessed and presented. Solid state media – typically flash – provides very fast access to data but can fail without warning. This is because it is typically subject to a limited number of program/erase cycles, as well as 'read/disturb' effects. Storage at scale also means the percentage likelihood of failure increases.		
<b>Additional Comments</b> Early generations of media and cheap giveaways are not robust and can deteriorate significantly over time.		


<b>Old or Non-current Offline Video Games</b> <p>Older or non-current video games designed and played on platforms and devices that are no longer supported online. These include video games that have a dependency on communication with servers that no longer exist (loss has already happened) and others where the dependencies are to do with old hardware/software where there is more potential to preserve the interaction.</p> <p>This entry also refers to offline gaming, understood here as the interactions and experiences of non-current offline video games that are playable on available on old hardware and software (console and PC technology). These remain playable due to a significant investment of skill and time by players but remain at risk with a lack of presence in safeguarded collections.</p>		
Group: Gaming	<b>Trend in 2021:</b>	
Added to List: 2017	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on a large group of people	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Older versions of games such as Civilization; EA Sports; Doom; Role-Playing Games (RPGs) such as The Elder Scrolls V: Skyrim. Games that rely on unique peripherals to play, including rhythm action titles such as Guitar Hero.		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Complex hardware dependencies or bespoke hardware; dependence on obsolete, low usage operating systems with no emulation pathway; complex intellectual property rights; use of older magnetic media; free distribution on magazines; loss of underlying code or gaming engine; limited or no commercial interest; dependency on remote servers that are closed.		
<b>'Endangered' in the Presence of Good Practice</b> emulation pathway; source code; trusted repository; large user community.		
<b>2021 Review</b> This entry was added in 2017 as a subset of 'Gaming,' It was rescoped in 2021 to bring in overlapping technical complexities with the 'Old or non-current video games' entry, which arise from preservation of software and hardware environments, and to narrow concerns specific to offline games and gaming. There is an active specialist market for older games which enables preservation but also skews it around commercial interests. The meaning of 'older' is open to interpretation, but any version more than ten years from release should be included here, even if and perhaps especially if there are more recent releases		
<b>Additional Comments</b> Work is happening in this space, and there is an avid amateur community working on this - how do we harness their work, so it is preserved rather than reinventing the wheel? There are a lot of older (online, typically MMPORG) games that are purely/mostly supported by its community fan base. For example, these communities (distinct to each game) create mods,		


updates, sometimes host servers, refine animations, etc., to keep these games playable and/or looking fresh for the rest of the community when it's no longer updated/supported by the game developer. What is interesting is that this is mainly for PC versions of the game because accessing the game on a computer just allows for more customization (Skyrim on PC vs. Skyrim on PlayStation).

<b>Old or Non-current Video Games</b> Older video games designed and played on platforms and devices that are no longer supported, especially those with complex but obsolete rights management or system dependencies. This group also includes older editions of games still published in newer editions		
Group: Gaming	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is now a priority	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> SimCity; Wolfenstein 3D; Lemmings; Animal Crossing		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Complex hardware dependencies or bespoke hardware; dependence on obsolete, low usage operating systems with no emulation pathway; complex intellectual property rights; use of older magnetic media; free distribution on magazines; loss of underlying code or gaming engine; limited or no commercial interest; dependency on remote servers that are closed.		
<b>'Endangered' in the Presence of Good Practice</b> Emulation pathway; source code; trusted repository; large user community;		
<b>2019 Review</b> This entry was added in 2019 as a subset of the previous 2017 entry for 'Gaming,' This entry for older games encourages greater consideration of the technical complexities which arise from preservation of software and hardware environments as well as the historic value that games have acquired. There is an active specialist market for older games which enables, preservation but skews it to commercial interests. The meaning of 'older' is open to interpretation, but certainly any version more than ten years from release should be included here, even if and perhaps especially if there are more recent releases		
<b>Additional Comments</b> It can be difficult to distinguish these from the related 'Offline gaming' entry, as some may be played not only offline in a self-contained way but also online or with online elements, which is why this entry is separated to note additional complexities in terms of systems dependencies that need to be considered. It is one thing to be able to get older software running on a different CPU architecture (emulation), but another thing entirely to recreate the peripherals (Wiimotes, plastic guitars, steering wheels, even bongo drums!) that are required to play the games as intended. There is a sort of kinaesthetic issue: one could perhaps use the accelerometers in modern smartphones to simulate the functionality of a Wiimote, but it will feel different in hand.  It might be possible to sub-divide between games that have a dependency on communication with servers that no longer exist (loss has already happened) and others where the dependencies are to do with old hardware/software where there is more potential to preserve the interaction.		

It is unclear who is responsible for preserving what, and games may be falling between the cracks. Often this is when the fan community steps in because developers will only update games that have real commercial benefit.




Open Source Intelligence Sources of Current Conflicts		
Open source social media and web content that supports crowd-sourced investigation and fact-checking to verify or refute claims of state agencies and rebel groups in the context of current political or military conflict.		
Group: Digital Legal Records	Trend in 2021:	Unanimous Decision
Added to List: 2019	No Change	Previously: Critically Endangered
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Social media sources relating to current conflicts, such as in Yemen or Syria.		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Loss of authenticity; lack of preservation agency; limited or no digital preservation capability.		
<b>‘Endangered’ in the Presence of Good Practice</b> Offline backup captured by the journalist or investigating authority;		
<b>2021 Review</b> This entry was added as a subset in 2019, as part of a broader ‘Open Source Intelligence Sources’ which the Jury split into three elements, relating to current, recent and historic sources. This entry relates in particular to materials relating to current and ongoing conflicts. Social media companies have a policy to take down or suppress content that they consider to be propaganda for terrorist groups. This has had the unintended consequence of deleting or suppressing content that was being used in open source investigation or fact-checking for journalistic or judicial purposes, and which may therefore be an impediment to refutation or prosecution. However, a new generation of cloud-based services, such as Hunchly, have emerged in the last few years, which allow investigators to copy and stabilize content to private accounts in the process of investigating it: so the ethical requirements of social media companies and the integrity of the investigation are both served. The Jury notes that such content remains at risk, and the process of investigation is slower than algorithmic deletion. Nonetheless, there is a notable difference in the investigation of current conflicts than historic ones where evidence has been lost.		
<b>Additional Comments</b> Important for social context and maybe picked up through inadvertently in other ways - but is ambiguous about who has ultimate responsibility for collecting and preserving this.  See: Higgins, E. (2019). <i>Bellingcat and beyond. The future for Bellingcat and online open source investigation</i> , iPres Conference 2019, Amsterdam, <a href="https://www.youtube.com/watch?v=kZAb7CVGmXM">https://www.youtube.com/watch?v=kZAb7CVGmXM</a> .		

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.		
Group: Sound & Vision	Trend in 2021:	Consensus Decision
Added to List: 2017	Trend towards reduced risk	Previously: Practically Extinct
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> AerTel; Electra; MetroText; Antiope-based systems; Ceefax; Teletext		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Lack of understanding; structure of information silos; Lack or loss of documentation; Uncertainty about intellectual property rights; Lack of funding or impetus		
<b>‘Endangered’ in the Presence of Good Practice</b> Captured within on-air broadcast recordings; active research and recovery programme		
<b>2021 Review</b> A number of important developments were reported since this entry was added in 2017, raising hope that collections can be recovered and re-used under certain circumstances. A trend towards decreased risk was noted in 2020, based on live-capture of broadcast output at the BBC and British Film Institute. Embedded signals can be assembled to access the Teletext content, and methods to recover such signals have been demonstrated (where the appropriate broadcast archive exist). In 2019, a proposal to add teletext graphics characters to the Universal Coded Character Set was accepted, making it easier to transmit and archive teletext and legacy computer graphics for archiving and preservation. Version 13.0 of the Unicode Standard, released in March 2020, included the addition of graphic characters that provide compatibility with early teletext broadcasting standards. The 2021 Jury agreed with the 2020 review and trend towards reduced risk. While it is not yet fully clear how such a signal could be made searchable or made available at scale, research has progressed with different threads of research and enthusiast and community led initiatives being brought together in the development of preservation tools and techniques. In light of these developments and active research and recovery efforts, the Jury supported a change from the <i>Practically Extinct</i> to <i>Critically Endangered</i> classification.		
<b>Additional Comments</b> I think this material is <i>Critically Endangered</i> and it will impact our understanding of pre WWW digital/electronic communications. When viewing impact in terms of impact on human life, it is minimal, but it does have an impact on understanding our history. The star quality is limited at this point because I do not think anyone is as concerned as internet historians or those involved in history and technology)		

From a cultural studies point of view, it is a huge loss: an important source of information about news and social mores of the time. Is there progress to report on this entry, meaning that elements of the problem have been resolved? Major national agencies have collections of off-air recorded television on videotape carriers, which are likely to contain the teletext data. Few, if any, have undertaken substantial extraction and preservation of the teletext in its own right, although many have digitized videotape carriers to digital file formats, which are now under preservation. An active Teletext enthusiast community has developed and has created programmatic solutions to the extraction of the teletext from the video files and emulation of the teletext display. A next step could be a collaboration between the official agencies and the enthusiast community to develop a systematic programme to extract teletext from off-air recordings for both preservation and access via emulation.

*See also:*

- *The Teletext Archive*, a depository for archive teletext services from around the world to be stored for research purposes, at <https://archive.teletextarchaeologist.org/>
- *Teefax*, the community teletext service, at <http://teastop.plus.com:8080/>
- *The Teletext Art Research Lab* website at <http://teletextart.co.uk/>
- *The Viewdata Frame Database*, that offers a browsable archive, at <https://db.viewdata.org.uk/>
- Unicode, Inc., (2020) *Announcing the Unicode Standard, Version 13.0*, <http://blog.unicode.org/2020/03/announcing-unicode-standard-version-130.html>

<b>Politically Sensitive Data</b> Digital content where the knowledge to preserve exists, and there is no threat to obsolescence, but where political interests may be served by elimination, falsification or concealment.		
Group: Political Data	<b>Trend in 2021:</b>	
Added to List: 2017	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Online News; social media and web-based campaigning; social media relating to 2016 UK/EU referendum; Promises made in Scottish independence referendum 2014; US Environmental Data; UK Public Finance Initiative (PFI) documents; Recordings of Leinster House;		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Opaque terms and conditions that facilitate deletion or obfuscation; lack of access to web-harvesting; significant lobby interest; change of administration; data resides in single jurisdiction; reputational risk to collecting institution		
<b>'Endangered' in the Presence of Good Practice</b> Robust political archives; robust preservation services for investigative journalists		
<b>2021 Review:</b> The nature and extent of political campaigning online continue to become more apparent. This has drawn attention to the manipulation of digital media but not explicitly the issue of deliberate deletion, alteration or concealment. GDPR provides a pretext for the disposal of records. The increased capability of archives to secure the content from outgoing governments and ministers is a source of encouragement, such as in Canada, accusations that the incoming Liberal government had wiped the memory of the outgoing Conservative government were shown to be unfounded. Nonetheless, there is a pressing need for a deep and comprehensive assessment of the risks faced by politically sensitive data and the impact which such deletions have on the public good. That another year should have passed without such an assessment is a matter of serious concern, leading to the 2020 trend towards increased risk, which the 2021 Jury agreed has continued with the continuation of significant political and economic upheaval, in part because of the pandemic, though also because of popular protest and the outcomes of elections around the world. Moreover, it has been widely reported that senior officials in government have avoided scrutiny and record-keeping laws by using self-deleting messaging applications. In these circumstances, politically sensitive records are likely to be at greater risk.		
<b>Additional Comments</b> This would be a break down in political systems and a break down in trust. I think time would be needed to advocate for the importance of preserving the records as well as immediate action. Maybe it is not the duty of archives or libraries to preserve the falsification but to preserve the constituent pieces to allow researchers to infer elimination, falsification or concealment. Web		


harvesting organizations will pick up web and possibly some social media. Will others pick up the rest?

The risk is high because some of the data will be in difficult to reach platforms. Some inevitable loss.

Is there a technological path to capturing and protecting this information before deletion or manipulation? This could well be the basis for crowdfunding and/ or crowd implementation.

*See also:*

- World Wide Web Foundation, The Open Data Barometer, which provides a global measure of how governments are publishing and using open data for accountability, innovation and social impact, which looks at the 30 governments that have adopted the Open Data Charter and those that, as G20 members, have committed to G20 Anti-Corruption Open Data Principles. Online at: <https://opendatabarometer.org/>
- Ovenden, R., (2020) Undelete our government. DPC Blog. Available at: <https://www.dpconline.org/blog/undelete-our-government>


<b>Records of Local Government</b> Records from local government (i.e., below the state level) which are required for transparency and may be in many diverse forms, but in which the local authority may lack the capacity to manage the complex digital preservation requirements that arise.		
Group: Public Records	<b>Trend in 2021:</b>	
Added to List: 2019	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to preserve materials in this group, requiring the application of proven tools and techniques.
<b>Examples</b> Born digital records of small and medium-sized agencies; fast-changing internal manuals, advice or policies shared electronically; records of care services; Documentation supporting long-lived contractual relations like Public Finance Initiatives; Organizational Slack channels; network drives; EDRMS; Email.		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Lack of preservation infrastructure; conflation of backup with preservation; loss of authenticity or integrity; Long-lived business processes; poor storage; churn of staff; significant volumes or diversity of data; poorly developed digitization; ill-informed records management; poorly developed migration or normalization; longstanding protocols or procedures that apply unsuitable paper processes to digital materials; encryption; political instability; lack of sustained funding.		
<b>'Endangered' in the Presence of Good Practice</b> Well managed data infrastructure; preservation enabled at the point of creation; carefully managed authenticity; use of persistent identifiers; finding aids; well managed records management processes; recognition of preservation requirements; strategic investment in digital preservation; preservation roadmap; participation in digital preservation community.		
<b>2021 Review</b> This entry was added in 2019 as a subset of a previous entry for 'Records of long duration from Local Government or Other Government Agencies.' The split was intended to allow greater concentration on the challenges that these distinct types of agency face. Local government typically operates across a broad range of digital formats and services, but it is unclear and unlikely that relatively small archival agencies are properly funded locally to support the wide range of digital preservation requirements that arise. The 2020 Jury noted the trend towards greater risk based on 2020 being a year of significant political and economic upheaval, putting additional strain on local government and its agencies. In these circumstances, already vulnerable records are likely to be at greater risk. The 2021 Jury similarly agrees with the continuation of this trend over the last year. They also noted examples like Grenfell to demonstrate the precarity of local government records, especially when these risks overlap with records of non-governmental agencies, resulting in significance and impact of loss, the impetus for action and call to governing frameworks where failing in enforcement (and depending on the jurisdiction).		
<b>Additional Comments</b>		

Significant research by the UK National Archives into Local Government Archives in England underlines the digital skills shortages that exist, especially with respect to preservation. There may be a benefit from splitting into a) legally required public record and b) additional information that may enrich our digital preservation of society. My assumption was that the roles and requirements for records management are clearly defined, but if this is not the case and there are inadequate resources to match the requirement, then the risk goes up.

*Case Studies or Examples:*

- The Grenfell Tower fire and Grenfell Tower Inquiry illustrate the precarity of local government records, especially when third-party contractors are involved. Not only does it show the potential impact of aggravating conditions for Records of Local Government, but it also applies to those of Records of Non-Governmental Agencies. See: <https://www.grenfelltowerinquiry.org.uk/>
- In Scotland, there is record keeping legislation that is relevant and governs some of this, such as the Public Records Scotland Act of 2011, online at <https://www.nrscotland.gov.uk/record-keeping/public-records-scotland-act-2011>
- The work and outputs of the EDRMS Preservation Taskforce, such as the EDRMS Preservation Toolkit, may be helpful for guidance in this context <https://www.dpconline.org/digipres/implement-digipres/edrms-preservation-toolkit>



<b>Records of Non-Governmental Agencies</b> Records of independent agencies and contractors that act on behalf of the state in the delivery of public services, and which may be present in many diverse forms, but for which the NGO or contractors may lack the capacity to meet the complex digital preservation requirements that arise, or may have a business motive to minimize or ignore requirements for the maintenance of the record		
Group: Public Records	<b>Trend in 2021:</b>	
Added to List: 2019	<b>Towards Greater Risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques..
<b>Examples:</b> Born digital records of small and medium-sized agencies; fast-changing internal manuals, advice or policies shared on intranets or EDRMS; records of care services; historic guidelines and manuals which evidence 'best practice at the time'; Documentation supporting long-lived contractual relations like Public Finance Initiatives; Organizational Slack channels; network drives; EDRMS; Email		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Lack of preservation infrastructure; conflation of backup with preservation; loss of authenticity or integrity; Long-lived business processes; poor storage; churn of staff; significant volumes or diversity of data; poorly developed digitization specifications; ill-informed records management; poorly developed migration or normalizations specifications; longstanding protocols or procedures that apply unsuitable paper processes to digital materials; encryption; political instability; lack of sustained funding; denial of responsibility; failure to include archives within contract from commissioning agency.		
<b>'Endangered' in the Presence of Good Practice</b> Well managed data infrastructure; preservation enabled at the point of creation; carefully managed authenticity; use of persistent identifiers; finding aids; well managed records management processes; application of records management standards; recognition of preservation requirements at highest levels; strategic investment in digital preservation; transfer protocols to public archive; participation in digital preservation community.		
<b>2021 Review:</b> This entry was added in 2019 as a subset of a previous entry for 'Records of long duration from Local Government or Other Government Agencies.' The split was intended to allow greater concentration on the challenges that these different types of agency face. Non-governmental organizations typically operates across a broad range of digital formats and services acting on behalf of public sector. The 2020 Jury noted the trend towards greater risk based on 2020 being a year of significant political and economic upheaval, putting additional strain on NGOs in these circumstances already vulnerable records are likely to be at greater risk. The 2021 Jury similarly agrees with the continuation of this trend over the last year. They additionally noted examples like Grenfell to demonstrate the precarity of records of non-governmental agencies, especially when these risks overlap with those of local government, resulting in significance and impact of		




loss, the impetus for action and call to governing frameworks where failing in enforcement for these agencies (e.g., examining current recordkeeping regimes keeping them accountable).

**Additional Comments:**

*Case Studies or Examples:*


- The Grenfell Tower fire and Grenfell Tower Inquiry illustrate the precarity of local government records, especially when third-party contractors are involved. Not only does it show the potential impact of aggravating conditions for Records of Local Government, but it also applies to those of Records of Non-Governmental Agencies. See: <https://www.grenfelltowerinquiry.org.uk/>
- There can be some grey areas depending on the legislative context. The Public Records Scotland Act 2011, for example, covers government agencies and any non-government org contracted to do work on behalf of government agencies, online at <https://www.nrscotland.gov.uk/record-keeping/public-records-scotland-act-2011>

<b>Records of Quasi Non-Governmental Agencies</b> Records from agencies at arms-length to government whether locally, nationally or internationally. They may be required to maintain archives for the purposes of transparency, sometimes for extended periods, and sometimes in diverse and complicated forms.		
Group: Public Records	<b>Trend in 2021:</b>	
Added to List: 2019	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is now a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Records of non-executive state or national agencies; museum or leisure trusts; industry or public regulators; public audit services; public-good funding and investment agencies; autonomous and semi-autonomous public agencies; sovereign wealth funds; public/private partnerships; publicly owned companies.		
<b>'Practically Extinct' in the Presence of Aggravating Conditions'</b> Lack of preservation infrastructure; conflation of backup with preservation; loss of authenticity or integrity; Long-lived business processes; poor storage; churn of staff; significant volumes or diversity of data; poorly developed digitization specifications; ill-informed records management; poorly developed migration or normalizations specifications; longstanding protocols or procedures that apply unsuitable paper processes to digital materials; encryption; political instability; lack of sustained funding.		
<b>'Endangered' in the Presence of Good Practice</b> Well managed data infrastructure; preservation enabled at the point of creation; carefully managed authenticity; use of persistent identifiers; finding aids; well managed records management processes; application of records management standards; recognition of preservation requirements at highest levels; strategic investment in digital preservation; preservation roadmap; participation in digital preservation community.		
<b>2021 Review</b> This entry was added in 2019 as a subset of a previous entry for 'Records of long duration from Local Government or Other Government Agencies.' The split was intended to allow greater concentration on the challenges that these different types of agency face. Records of quasi non-governmental agencies are at arm's length to government, but the 'QuaNGO' or 'ALEO' (Arms-Length Executive Organization) may lack the capacity to meet complex digital preservation requirements that arise, nor be able to deposit in the government archive. The 2021 Jury added that arm's length bodies are still public bodies, and they have a duty of care for maintaining evidence of their actions and transactions. They often receive public funding, and depending on the archives, legislation may be required to transfer to an archive. The issue is when there is a lack of clarity regarding the recordkeeping requirements or neglect of records and information once it has outlived its usefulness. These bodies still create records that affect citizen lives and have a		

duty to document, and therefore the *Critically Endangered* classification is better suited. There is also a 2021 trend towards greater risk when looking at the call for clarity and action to reduce the impact of loss to citizens during a period of significant political and economic upheaval when there is additional strain on agencies.

**Additional Comments**

Although the split draws attention to the different pressures faced by QuaNGOs it could be further subdivided into legally required public records and additional information that may enrich our digital preservation of society. The classification assumes that the roles and requirements for records management are clearly defined, but if this is not the case or there are inadequate resources to match the requirements, then the risk goes up.

<b>Smart Phone Apps</b> Apps created for smartphones. Many are deprecated quickly but others survive through multiple update cycles. It is hard to maintain version control and often dependent upon the company that publishes them. There is no clear agency or mandate to record or collect.		
Group: Apps	Trend in 2021:	
Added to List: 2017	<b>No Change</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> London 2012 app; BBC Olympic app; Apps published for Apple iOS 10 or earlier		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Device dependence; poor documentation; uncertainty over IPR; short term contracts; lack of skills, commitment or policy from corporate owners; rapid churn of OS; shifting business requirements of app resellers; dependence on exotic or obsolete formats or OS processes;		
<b>'Endangered' in the Presence of Good Practice</b> Strong documentation; version control for code and compiled app; emulation enabled; designated repository taking preservation responsibility and capacity to deliver		
<b>2021 Review</b> This entry was added in 2017 to draw attention to the challenges of software preservation and the extraordinary velocity of the market for apps. Given the speed of change, it is hard to see how digital preservation efforts can keep pace. The 2019 Jury noted that splitting this entry into different groups based on the platform would clarify differences, although though the risks would be largely the same. The 2021 Jury discussed this further in light of the new 2021 'Smart Phone Gaming' entry, which can be considered a subset of this category as well as Gaming in which it is currently assigned. This is, therefore, an entry worth revisiting in the next review. The Jury also discussed the impact of loss; some loss seems inevitable due to the changes that have taken place in mobile ecosystems over the past 15 years, having already rendered old apps and versions of apps inaccessible, and how important is it to preserve every update? When discussing whether there is any 2021 trend, the jury discussed covid check-in apps as examples but similarly questioned the impact of this loss (do we want to preserve every version who if in the future access to the documentation to re-create the app for a particular purpose?). The consensus was there was no significant trend towards greater or reduced risk over the last year, but this may warrant further discussion in the next review.		
<b>Additional Comments</b> Old versions of apps are completely lost to most users: once you upgrade an app, you typically cannot go back. Perhaps iOS is more critical - at least with Android, you can often get .apk from the internet separate from the marketplace. The NSRL contains hundreds of thousands of mobile applications which are not being actively preserved but could be if a mandate existed. An extension to Legal Deposit might be possible.		


The faster we act, the less we will lose. It is unlikely that there will ever be one agent with a mandate to collect and different apps available in different countries, so a network of national organizations would be needed. The companies that create these apps are the key to the licensing challenges, and conversation with them is necessary, though it would need to happen immediately in order to negotiate the right to preserve/escrow both apps, operating systems, documentation, and phone development emulators.

Apps often provide a secondary view of the primary data, and in many cases, that data is more important and arguably easier to preserve. So, although this is an eye-catching entry, it is not as significant as others.

It depends on what apps we are speaking of. I would say we've seen the impact of ministers deleting content on their phones, some of which will be held by apps. I think a major risk here is messaging apps like WhatsApp because they capture decisions that are easily deleted, compromising accountability or transparency.

*See also:*


- Siddique, H. (2021) *UK government admits ministers can use self-deleting messages*. The Guardian. 13 Jun, <https://www.theguardian.com/politics/2021/jun/13/uk-government-admits-ministers-can-use-self-deleting-messages>
- Siddique, H. (2021) *Cabinet policy obliges ministers to delete instant messages*. The Guardian. 12 Oct, <https://www.theguardian.com/politics/2021/oct/12/cabinet-policy-ministers-delete-whatsapp-messages>

<b>Smart Phone Gaming</b> Smart phone gaming is an example of contemporary digital culture but is often considered less important than other games. Many require community engagement for a game to function, similar to Massively Multiplayer Online (MMO) Gaming.		
Group: Gaming	New Entry	
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment in one year.	<b>Significance and Impact</b> The loss of tools, data or services within this group would impact on a large group of people and sectors.	<b>Effort to Preserve / Inevitability</b> It would require a major effort to prevent or reduce losses in this group, including the development of new preservation tools or techniques.
<b>Examples</b> Pokémon GO, Candy Crush, FG/O, Clash of Clans, Angry Birds		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> uncertainty over IPR; short term contracts; lack of skills, commitment or policy from corporate owners; rapid churn of OS and updates; shifting business requirements of app resellers; dependence on exotic or obsolete formats or OS processes; loss of underlying code or gaming engine; limited or no commercial interest; dependency on remote servers that are closed; limited recognition of value of game play; over-dependence on goodwill of ad-hoc community;		
<b>'Endangered' in the Presence of Good Practice</b> IPR supportive of preservation; strong documentation; version control for code and compiled app; source code; emulation pathway; trusted designated repository or community taking preservation responsibility and capacity to deliver; inclusion by agencies that collect games on other platforms;		
<b>2021 Review</b> This is a new entry submitted through the 2021 open nomination process. There are overlaps between this entry and others relating to both Gaming and Smart Phone Apps. The 2021 Jury recruited additional expertise for a recommendation on which category it better fits and agreed with the expert recommendation to assign the Gaming category but keep as a separate entry to emphasize that smart phone gaming shares preservation issues with video games that are exacerbated by issues unique to smartphones. Smart phone gaming is an example of contemporary digital culture that is often not considered as important as other games and can require community engagement for a game to function similarly to MMOs. While it shares the challenges of mobile gaming (larger volume of titles compared to consoles), the entry was limited to smart phone gaming for two reasons: 1. in relation to gaming, mobile gaming does not have the same community of support for DP and therefore is a special case, and 2. device dependence and lack of infrastructure to recreate digital environment via smartphone.		
<b>Additional Comments</b> Smart phone games such as Pokémon GO, etc. has large active communities of players and fandoms. The argument is that in relation to gaming, mobile gaming does not have the same community of support for DP and therefore is a special case. No clear agency with the intent to collect smartphone games, though there are agencies that collect games on other platforms.		

It is unlikely that there will ever be one agent with a mandate to collect and different smart phone gaming apps available in different countries, so a network of national organizations would be needed. The companies that create these are the key to the licensing challenges, and conversation with them is necessary, though it would need to happen immediately in order to negotiate the right to preserve

*Case Studies or Examples:*

- Example of Flappy Birds game being acquired by the V&A, see Volsing, K (2014) *The Rise and Fall of Flappy Bird and the collecting of the V&A's first App* • V&A Blog. V&A Blog. Available at: <https://www.vam.ac.uk/blog/news/the-rise-and-fall-of-flappy-bird-and-the-collecting-of-the-vas-first-app>
- There is a Wikipedia entry providing a list of the most played mobile games by player count. See: [https://en.wikipedia.org/wiki/List\\_of\\_most-played\\_mobile\\_games\\_by\\_player\\_count](https://en.wikipedia.org/wiki/List_of_most-played_mobile_games_by_player_count)

Supporting Digital Materials for Museums and Galleries		
All manner of digital materials held in museums to support the access, interpretation and management of physical collections, including material supporting collections management, conservation data, interpretative materials, oral history, and generated through the course of work and research. It excludes digital material formally accessioned into museum collections.		
Group: Museum Data	Trend in 2021:	Consensus Decision
Added to List: 2019	Trend towards greater risk	Previously: Endangered
<b>Imminence of Action</b> Action is recommended within three years, detailed assessment within one year.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a small effort to address losses in this group, requiring the application of proven preservation tools or techniques.
<b>Examples:</b> Digital photography, video, sound, oral histories, collections management documentation, collections management systems, conservation records or data, records of exhibitions, interpretive materials, records, archives, operational or business records, research outputs, correspondence, 3d digitization.		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Lack of repository infrastructure; external dependencies; dependency on cloud provider; poor storage; churn of staff; significant volumes or diversity of data; poorly developed digitization specifications; ill-informed records management policy; conflation of access with preservation; poorly developed migration or normalization; longstanding protocols or procedures that apply unsuitable paper processes to digital materials; inability to assume responsibility; encryption		
<b>‘Endangered’ in the Presence of Good Practice</b> Repository and preservation infrastructure; strategic leadership; well-developed digitization and migration pathways; participation in the global digital preservation community; Business continuity records, i.e., accession records, catalogues loss should be avoidable with good practice		
<b>2021 Review</b> This entry was added in 2019 under ‘Digital Materials in Museums and Galleries’ to bring together submissions from the open nomination process. The 2019 Jury took the view that museums make creative use of digital technology but have weak incentives to preserve the outputs when compared to libraries or archives. Museums take responsibility for the care of physical collections in the long term, which brings a derived requirement to preserve digital materials relating to those collections. The museum sector lacks preservation capability. The 2020 Jury noted that the entry was very broadly defined, and whilst it is useful to draw attention to the challenges, it might usefully be disaggregated in the future. The 2021 Jury agreed, commenting on a potential misunderstanding with the entry based on what examples first come to readers’ minds when reading the description; some interpreted the material as databases and catalogues, others interpreted it more as business continuity records and interpretive outputs through websites, and readers of the BitList may also have other interpretations with the broad ‘Digital Materials in Museums and Galleries’ title. While these all may fit under the purposively broad entry, the Jury voted to rename and rescope the entry with emphasis on supporting material for collections management, conservation data, interpretative materials, and oral history.		



The 2020 Jury also noted a trend towards greater risk based on many museums and galleries, which often rely on visitors for income, closed for extended periods. In those circumstances of economic dislocation digital materials in museums and galleries records are likely to be at a greater risk than in 2019. This trend continues for 2021 with the major funding crisis in museums after Covid 19.

**Additional Comments**


Much of these records already exist in hybrid and unorganized forms. Most museums do not have a grasp on the scale of their problems (especially big ones that have been doing digital forms of collections management for years and have lots of floppy discs, CDs and legacy systems that have not been properly migrated). Not convinced that museums always have a handle on their data management and apply the appropriate conservation methods and practices to this material to ensure its perennity. Moreover, this sector is incredibly diverse and different needs are likely to exist in different kinds of museums or galleries.


It may be useful to disaggregate this entry again in the next review, pulling out born digital material and content that is potentially at greater risk to create new subsets as well as separating those with less risk. For example:

Databases and catalogues may have lower risk as they are more likely to be preserved because they are periodically updated. Museums have significant incentives to maintain collection catalogues for audit purposes and to share these with the public. Such databases are likely to face less severe preservation challenges than data generated about the collection through museum operations, conservation or research. But there can still be a lack of preserving the content in databases and a long way to go in smaller institutions (yes, tech and vendors improving, but the issue is with institutions).

Interpretive outputs through websites may have reduced risk as web outputs if captured as part of national web archiving initiatives. Exhibition catalogues and interpretation of collections are often published online in research papers, etc.

Related data sets present different challenges too. For example, 3D digitization is a particular challenge for museums as preservation standards are deficient and easily overlooked in the enthusiasm for fashionable or eye-catching trends which purport to extend access.

Unpublished Research Data from Government Researchers		
Data sets and research outputs produced in the course of government research but never shared or made available outside of the initial research. In particular, the risk classification applies to research data under government embargo, restrictions due to sensitivities, classification issues, and/or materials suppressed for ideological reasons.		
Group: Research Outputs	Trend in 2021:	Consensus Decision
Added to List: 2019	Towards Greater Risk	Previously: Critically Endangered
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> It would require a major effort to prevent or reduce losses in this group, possibly requiring the development of new preservation tools or techniques.
<b>Examples</b> Data sets or research outputs produced for agencies that have closed or have had funding withdrawn from research initiatives, research data from government agencies that is no longer active		
<b>‘Practically Extinct’ in the Presence of Aggravating Conditions</b> Lack of access to archival services; sudden or un-anticipated closure; loss of implicit knowledge from destabilized or demoralized staff; encryption		
<b>‘Endangered’ in the Presence of Good Practice</b> Archival responsibility well developed; documentation; published through research channels.		
<b>2021 Review</b> This entry was added in 2019 under ‘Unpublished Research Data from US Govt Researchers’ It has significant overlaps with other entries in the research outputs group but was retained to draw attention to two realities: firstly that research outputs are not simply a matter for academic institutions, and that government is, in fact, a major producer of research data; and secondly that political instability and threats to the continuity of government services are a significant preservation risk. The 2019 entry description noted that while it related to the US, it did not mean that other jurisdictions are immune from political instability, and comment by the 2019 Jury that politically inconvenient research outputs face particular and immediate threats of which the digital preservation community should be cognizant. The 2021 Jury agreed with this concern and broader applicability but recommended that this should be more explicit, and both title and description should be changed to broaden and include governments across national and international contexts. This change does mean that the risk profile will range and depend on the political system, the political change and the measures in place to save and reuse data from disbanded research projects; in other words, there may be instances where the unpublished research data in one country may fall under the <i>Vulnerable</i> category.		
<b>Additional Comments</b> The US made the news as part of the last government, but this is probably an issue in other countries as well and is, therefore, a category that could be made more generic. One question to ask is whether the research data is considered of long-term value or considered ephemeral?		


<b>Web Domains with No Legal Deposit</b>		
Web archiving is a recognized specialism within digital preservation, able to capture large quantities of material with routine and standards-based tools. But there are significant issues of intellectual property rights associated with website capture and republication. In many jurisdictions, but by no means all, those obstacles are overcome by regulations that enable a national library or other 'legal deposit' agency to copy and preserve content. Where no such permission exists, there is a significant risk of loss.		
Group: Web	<b>Trend in 2021:</b>	Consensus Decision
Added to List: 2019	<b>Towards Greater Risk</b>	Previously: Critically Endangered
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on many people and sectors.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Domains registered without a country code; domains with a country code but weak or unenforceable legal deposit permission to harvest.		
<b>'Practically Extinct' in the Presence of Aggravating Conditions</b> Rapid churn of websites; lack of access to Internet Archive harvest; contentious content; encryption; digital rights management; non-standard content management.		
<b>'Endangered' in the Presence of Good Practice</b> Permissive approach to Legal deposit.		
<b>2021 Review</b> This entry was added in 2019. It is characterized by regulatory barriers rather than technical ones, though the pace of change in web technologies, as well as the growth of web content, mean that significant technical challenges still exist. The 2019 Jury noted that local conditions are also a significant factor. For example, websites often also fall under public records legislation or are important elements of corporate records: and so important parts of the web are harvested even when there is no explicit legal deposit legislation. Moreover, the Jury particularly recognizes the work of the Internet Archive to capture and preserve content. Even so, there are significant gaps in web archiving, and in too many cases, it is regulation that is the barrier. The 2021 Jury agreed with this description and classification but added that in some limited instances, pywb tools (as opposed to automated web crawlers like Heritrix) could effectively capture the look and feel of a platform interface, preserving legacy versions for users to interact with in the future. However, pywb tools are manual and therefore cannot address the scale of this issue. They also do not capture interfaces in a way that makes it possible to recreate them in the future, only interact with a defined set of web pages. For this growing issue of scale, the 2021 Trend is towards greater risk.		
<b>Additional Comments</b> Unless the Internet Archive is picking these up, the early web or permission regimes are in place, and these early instances are gone forever and will continue to be lost.		


## 11. Practically Extinct



Digital materials are listed as *Practically Extinct* when examples cannot be identified or are inaccessible by most practical means and methods. It does not assume that the material is lost, but rather that loss is imminent and immediate action is required to avoid loss. It includes material where recovery is possible in very small samples but is impractical or has not been demonstrated at scale.

This classification includes *Critically Endangered* materials in the presence of aggravating conditions.

<b>Adobe Flash Animations and Interactive Applets</b> Animations, games, and other interactive applets created with Macromedia Adobe Flash Player and Shockwave Flash, along with their accompanying websites. These are primarily .swf files, but they can also include networked collections of .swf files and external assets, as well as the web pages where they are displayed.		
Group: Web	<b>Trend in 2021:</b>	
Added to List: New Entry	<b>Trend towards greater risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> It would require a major effort to prevent losses in this group, such as the development of new preservation tools or techniques.
<b>Examples</b> Flash and Shockwave based games; cartoons; interactives		
<b>'Critically Endangered' in the Presence of Good Practice</b> Migration plan initiated; supported in multiple browsers; security vetted; emulation pathway		
<b>2021 Review</b> This entry was added in 2019. The 2019 jury noted that flash animations and applets were a mainstay of interactive web design from the late 1990s. Flash animations and interactives are created using tools supplied by the Adobe of the same name. Although Flash enables the development of sophisticated interaction at low cost over the web, it has had a chequered history in terms of browser support and has been plagued by security concerns. The 2020 Jury added the trend towards greater risk based on the indication by Adobe for some time that there would be withdrawal of support to Flash Animation. As noted by the 2021 Jury, the discontinuation and withdrawal of support did indeed occur. Flash is no longer supported, and loss has already occurred with Adobe's deprecation of Flash and lack of support in modern web browsers. For this reason, the classification has moved to <i>Practically Extinct</i> with a trend towards greater risk given the loss of dependence on Flash, which has resulted in new aggravating conditions for migration pathway, emulation pathway, source code, lack of capacity or motivation to support, no commercial interest.		
<b>Additional Comments</b> Flash represents a significant amount of the creativity of websites in the early 2000s including net-based art and cartoons. With extinction, archives will need to consider if it is possible to preserve interaction through the development of new web archiving and emulation tools and techniques		

Legacy Interfaces and Services Offered Online by Major Companies		
Online services with unique interfaces that change regularly and through those changes provide a different experience AND different content to their users.		
Group: Social Media	Trend in 2021:	Unanimous Decision
Added to List: 2019	No Change	Previously: Practically Extinct
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Interfaces to Gmail, Facebook, Google Docs, Hotmail, Ask Jeeves, Tweetdeck, TurboTax, MySpace, Quicken Online, and many others		
<b>‘Critically Endangered’ in the Presence of Good Practice</b> Robust and extensive web archives with strong documentation of search algorithms, ranking and personalization of interfaces.		
<b>2021 Review</b> This entry was added in 2019. The Jury noted that while there are overlaps with several other entries around social media and the web which pertain to content, this entry highlights the configuration of interfaces and, therefore, the ever-changing arrangement and presentation of content. Personalization means that the same query can produce quite different results to different users at the same time; the application of machine learning to behavioural surplus means the same may obtain different results at different points in time. That is over and above the rapid churn in the appearance of web interfaces. There is little appreciation of the implications for the use of online services and the potential for manipulations that arise. Moreover, the digital preservation community, which is historically concerned with data rather than interface, has only rudimentary tools to address this challenge. The 2021 Jury agrees with this description as well as the 2020 Jury’s conclusion of no trend towards greater or reduced risk but adds that the lack of standardized regulation and lack of transparency of operations by some of these platforms can still be problematic and make the content hosted there vulnerable to loss. For example, some content creators on YouTube may lose access to their content and accounts due to copyright infringement claims or reports of inappropriate content, which may or may not be supportable. The risk of loss is higher if the content is not stored anywhere else. Though some mitigation methods are available through the platform, this issue may only affect a small number of accounts.		
<b>Additional Comments</b> Some of the content/iterations of these are likely preserved to an extent within existing web archives but not as targeted collection efforts. As we've seen with myspace and other platforms where the platform producers decide to remove content or shut down rather quickly, it can be too late if this content has not been preserved already.		

Upgrading is compulsory - if really considered a problem, could an emulator be developed/used? This is like how some sites respond differently depending on which browser you are using - what is the significance or value in capturing all these differing user experiences?


Why can we see how online services behaved five years ago? Moreover, why cannot we see how they manipulated data to present content differently from how they now do such that the content we can access via them is different?

The authenticity of displaying social media content from 2014 through modern interfaces is questionable, and without recording the interface at the time, it is not currently possible to recreate older interfaces. You'd think the platform owners would have the older versions saved, but these are not available at the moment, and it would be worth engaging in a conversation about making them available to cultural heritage institutions for display purposes.

How far do we take this? The returns are likely to diminish. Who is taking responsibility to preserve? What are the platform creators doing to preserve this cultural history?

Some of this information is almost certainly lost already (some through deliberate erasure). The imminence of action depends on the type of institution.




<b>Non-standard Public Records</b> Records created in the course of public administration and subject to public records legislation but created on unofficial channels and platforms and therefore subject to unlawful destruction whether by accident or design.		
Group: Digital Legal Records	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Unanimous Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Content and messages from cloud-based instant messaging services (such as WhatsApp or Snapchat) that pertain to public administration and are subject to public records legislation but concealed from or inaccessible to archival agencies.		
<b>'Critically Endangered' in the Presence of Good Practice</b> Archival pathway; public officials briefed on the nature of public records and the penalties for illegal disposal; boundary between public and private correspondence; cloud services administered transparently; export functions.		
<b>2021 Review</b> This entry was added in 2019 as a subset of an entry in 2018 for 'Digital Legal Records and Evidence,' which the Jury split into four different entries in order to draw attention to the different challenges and priorities that arise. The 2019 Jury gave this entry the strongest indication of risk available. This group includes those records which may contain politically damaging or uncomfortable realities and thus be at risk of deletion and may be concealed from archival agencies whether by accident or design. The 2019 Jury also noted that the destruction of certain classes of public records is unlawful, whether or not it is deliberate. The 2020 Jury added the trend towards greater risk based on the 'pivot to digital' necessitated by the Pandemic resulting in widespread changes in workflow and in the platform for the delivery of government, with significant amounts of remote working. These changes happened rapidly, often without time to consider the preservation and record keeping implications. In those circumstances, it was reasonable to suppose the risks expanded in size as well as scope. The 2021 Jury agrees with this assessment and classification but found no significant increase or decrease to the 2020 trend. The 2021 Jury also adds the comment that there should be a balance between trying to preserve what has already been created using these channels and trying to educate against/prevent records from being created this way in the future.		
<b>Additional Comments</b> This is a 'small effort to fix' in terms of the technology to export data. But loss seems likely unless there is stronger monitoring and enforcement of the policy around this.		





Agencies responsible for the public record will not be able to completely control their public servants' use of unofficial channels (but could tighten), so they need methods to obtain from unofficial channels. Very important for public accountability and transparency of the state.

Obviously, these records should not be created using these channels in the first place. It is probably unlawful for those in public office. However, we know there has and always will be this kind of backdoor activity and pretty much always a scandal when it is revealed. The challenge is managing to collect it

<b>Older Open Source Intelligence Sources</b> Older open source social media and web content that supports crowd-sourced investigation and fact-checking to verify or refute claims of state agencies and rebel groups in the context of historic political or military conflict.		
Group: Digital Legal Records	Trend in 2021:	
Added to List: 2019	<b>No Change</b>	Consensus Decision
<b>Imminence of Action</b> Immediate action necessary. Where detected they should be stabilized and reported as a matter of urgency	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Social media sources relating to the Arab spring		
<b>'Critically Endangered' in the Presence of Good Practice</b> Offline backup documented and available for recovery;		
<b>2021 Review</b> This entry was added in 2019 from a nominated entry that was split into three subsets by the 2019 Jury relating to current, recent, and historic sources. This entry relates in particular to materials published at the time of the 'Arab' spring. Social media companies had initially taken little or no action with respect to social media content in conflict zones, taking the view either that they were mere technical platforms and therefore not responsible for editorial; or that the platforms were being used largely for social good, loosening the control of the media from oppressive regimes. However, as the Arab Spring progressed, the companies came under significant pressure to monitor content with more care, in part because terrorist groups had begun using social media platforms for propaganda purposes. The social media companies responded by implementing algorithms that removed or deleted content. This had the unintended consequence of deleting or suppressing content that was being used in open source investigation for journalistic or judicial purposes and may have resulted in refutation or prosecution. The 2019 Jury recognized the duty of care that social media companies have towards their users and is in no sense seeking to have that material re-published on the open web but noted the unintended consequence for journalists and investigatory authorities from the rush to deletion, illustrating how this entry further underlines the relative fragility of all social media content. The 2021 Jury agrees with the current classification and description.		
<b>Additional Comments</b> This is important for social context but may be picked up inadvertently through other ways; it remains ambiguous about who has ultimate responsibility for collecting and preserving this.  See also: <ul style="list-style-type: none"> <li>Al Sayyad, N. and Guvenc, M., (2015) <i>Virtual Uprisings: On the Interaction of New Social Media, Traditional Media Coverage and Urban Space during the 'Arab Spring.'</i> Urban Studies, <a href="https://doi.org/10.1177/0042098013505881">https://doi.org/10.1177/0042098013505881</a></li> </ul>		

- Dougherty, R., (2021) Documenting Revolution in the Middle East. [online] Focus on Global Resources, CRL. <https://www.crl.edu/focus/article/7437>
- Higgins, E. (2019). *Bellingcat and beyond. The future for Bellingcat and online open source investigation*, iPres Conference 2019, Amsterdam, <https://www.youtube.com/watch?v=kZAb7CVGmXM>.

Pre-WWW Videotex Data Services and Bulletin Board Services		
Pre WWW telephone and television information services that allowed a degree of user interaction and data retrieval with modem-based two way communication.		
Group: Sound & Vision	Trend in 2021:	Unanimous Decision
Added to List: 2017	No Change	Previously: Practically Extinct
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Prestel, Minitel, VidiTel and Videotex NL, Alex, BelTel, FidoNet		
<b>'Critically Endangered' in the Presence of Good Practice</b> Offline backup documented and available for recovery;		
<b>2021 Review</b> This entry was added in 2017 and there has been no evidence to document any change in the initial classification that such data is <i>Practically Extinct</i> . There may be examples residing in offline backups of services taken at the time, but these are likely to have deteriorated rapidly. Therefore the Jury calls on anyone with such collections to act quickly to stabilize and recover content. The 2021 Jury agrees with the current classification and description with no change to trend.		
<b>Additional Comments</b> Is there any real value with this type of content apart from nostalgia and understanding the evolution of technology? I think that this material is <i>Critically Endangered</i> and it will impact our understanding of pre WWW digital/electronic communications. When ranking 'impact,' I considered the loss in terms of impact on human life, which is minimal, but it does have an impact on understanding our history. From a cultural studies point of view, it is a huge loss.  This is also something that links to community archives and community heritage - early online forums were a place of community development and community creation.  Presumably, a database is sitting behind these types of things and has emulation been tried? Although there is no structured collection of this material, many individuals have archives, and a campaign of the nature of 'Missing believed Wiped' might be effective Almost impossible to get this data back (we can hope for some disks to show up one day that has traces on them).		

<b>Unpublished Research Data</b> Data sets produced in the course of research but never shared or made available outside of the initial research team		
Group: Research Outputs	Trend in 2021:	
Added to List: 2019	<b>Trend towards reduced risk</b>	Consensus Decision
<b>Imminence of Action</b> Action is recommended within twelve months, detailed assessment is a priority.	<b>Significance of Loss</b> The loss of tools, data or services within this group would impact on people and sectors around the world.	<b>Effort to Preserve</b> Loss seems likely: by the time tools or techniques have been developed the material will likely have been lost.
<b>Examples</b> Unpublished research data		
<b>'Critically Endangered' in the Presence of Good Practice</b> Replication and documentation; data management plan; preservation pathway agreed		
<b>2021 Review</b> This entry was added in 2019 as a subset of the 'Unpublished Research Outputs' reported in 2018, which was split into entries to draw attention to the different preservation requirements and concerns that arise. This entry relates specifically to research data which has not been shared or published by any means and is thus in contravention of the 'FAIR' principles which require data to be Findable, Accessible, Interoperable and Reusable. Without proper planning, research data can have a high barrier to re-use, especially where documentation is lacking. The Jury takes the view that documentation and re-use go hand in hand, and researchers should be under no illusions that data not documented or shared faces material and immediate risks of extinction. Over the years, there have been numerous attempts to address the risk of data loss, and it was the 2019 Jury's hope that this is now a small group. The 2021 Jury agreed with the description and classification and did see that there has been a trend towards reduced risk in light of more robust collaborative initiatives to jointly address the risk of data loss in and across research communities.		
<b>Additional Comments</b> If we do not know it exists, does it exist? It may also be that in certain circumstances this includes data that is unfavourable and has intentionally not been published.  If perceived as high-value, someone in the research team will likely take steps to ensure it is protected. We can be proactive and offer advice, but ultimately it is down to them. We cannot keep everything!  This is a wide field, so the scale and impact are hard to describe, but the risk is higher than papers due to potential file format complexity.  Success is dependent on how successful an institution's research data management communications are. Advocacy and research are needed to show the scale of the problem, as well as education regarding open science and preservation.		

## 12. Appendix

### Scoring Criteria

Classification	<i>How should we categorize the risks here?</i>	
5	<b><i>Practically Extinct</i></b>	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 Critically Endangered materials in the presence of aggravating conditions.
4	<b><i>Critically Endangered</i></b>	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 Endangered materials in the presence of aggravating conditions.
3	<b><i>Endangered</i></b>	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 Vulnerable materials in the presence of aggravating conditions.
2	<b><i>Vulnerable</i></b>	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.
1	<b><i>Lower Risk</i></b>	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.

Significance	<i>What is the significance, value and/or 'uniqueness' of the digital material?</i>
10	Unique, irreplaceable, invaluable, of global interest and importance to a large group of people
9	Unique, irreplaceable, invaluable, of international interest and importance to a large group of people
8	Unique, irreplaceable, invaluable, of national interest and importance to a large group of people
7	Unique, irreplaceable, invaluable, of community/local interest and importance to a large group of people
6	Irreplaceable, invaluable, of interest and importance to a large group of people
5	Invaluable and significant to a smaller group of people
4	Of some value and significance to a smaller group of people
3	Of some value or significance to a smaller group of people
2	Derivative and easily reproducible
1	Limited significance

Impact	<i>What is the extent to which the loss of the digital material would be felt?</i>
10	Globally traumatic, destabilizing impact to a large group of people
9	Internationally traumatic, destabilizing impact to a large group of people
8	Nationally traumatic, destabilizing impact to a large group of people
7	Locally traumatic, destabilizing impact to a large group of people
6	Impact to a large group of people
5	Impact to a smaller group of people
4	Inconvenient
3	Inconvenient, but quickly recoverable
2	Inconvenient to a smaller group of people, but quickly recoverable
1	Briefly inconvenient

<b>Inevitability of Loss (Effort to Preserve)</b>	<i>Can it reasonably be avoided?</i>
<b>5</b>	Inevitable
<b>4</b>	Loss seems likely
<b>3</b>	Major effort to fix
<b>2</b>	Small effort to fix
<b>1</b>	Entirely avoidable

<b>Imminence of Action</b>	<i>If you had one of these, how quickly would you need to act to save it?</i>
<b>5</b>	Act immediately within 6 months
<b>4</b>	Around 12 months
<b>3</b>	Less than 3 years
<b>2</b>	Within 5 years
<b>1</b>	More than 5 years

## Jury Comments on Scoring Criteria and Metrics

- Significance and impact depend on the subject matter included on the portable media.
- When ranking 'impact,' I considered the loss in terms of impact on human life, which is minimal, but it does have an impact on understanding our history
- I find the emphasis on 'trauma' in your Impact rubric problematic. I suggest that it needs finessing. Currently, it implies that trauma is the only way that loss matters, and that will skew ratings towards those records whose loss will be traumatic. As such, I have ranked many of these on the basis that the impact would be to a large number of people, often having international significance but not 'traumatic' as such. The 'trauma' designation seems overdetermined, and few digital objects would meet the threshold of causing trauma to a large group of people.
- The inevitability of loss depends on the type of media and file types on the media.
- I've also struggled with the options for the 'inevitability of loss' metric, as the answer choices seem to me to be asking two distinct things - the likelihood of loss and the effort required to prevent loss. In my experience/opinion, these are not mutually exclusive. I can think of examples in which 2 of the answer choices can co-exist, such as with the web entries: loss is inevitable with legacy research web collections, but also this can be fixed.
- For imminence of action it will all depend on the format of the records. Correspondence or photographs may be left for longer, but recordings will need closer attention, especially if it is a bespoke recording format
- I want to contextualize my scores with the caveat that not all defunct platforms present the same risk, significance, inevitability of loss, or imminence of action. Similarly, the value of the content lost will vary from platform-to-platform and user-to-user. In many cases, the significance will be personal - loss of correspondence or interaction with a loved one or loss of professional work stored exclusively on the platform
- I found it challenging to use the same metrics for evaluating content-based preservation concerns (records of local government, etc.) compared to carrier-based concerns (hard disks). I'm not sure how to assign meaningful scores to the loss of hard drives. All hard drives? Some hard drives? Doesn't it depend on what content is stored on them? As someone brought up in the meeting, there's a distinction between something being at risk of loss vs. worthy of preservation efforts.

The Digital Preservation Coalition  
11 University Gardens,  
Glasgow G12 8QQ  
Scotland

[www.dpconline.org](http://www.dpconline.org)  
[Info@dpconline.org](mailto:Info@dpconline.org)

The Digital Preservation Coalition (DPC)  
exists to secure our digital legacy.

We enable our 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. We achieve our aims through advocacy, community engagement, workforce development, capacity-building, good practice and good governance.

© Digital Preservation Coalition 2021