PRESERVING COMPLEX DIGITAL OBJECTS

Workshop

Caylin Smith
British Library & UK LDLs
United Kingdom
caylin.smith@bl.uk
0000-0001-6340-5708

Patricia Falcao
Tate
United Kingdom
Patricia.Falcao@tate.org.uk
0000-0003-2798-5631

Sara Day Thomson
Digital Preservation Coalition
United Kingdom
sara.thomson@dpconline.org
0000-0002-3896-3414

Abstract - This workshop addresses the opportunities and challenges generated by complex digital objects - objects created using innovative technologies - and current strategies for preserving them. The workshop will present a definition of complex digital objects with examples and an overview of the preservation challenges they pose, such as deciding what to collect and problem-solving for software and hardware dependencies. Using case studies, participants will identify and analyze a set of challenges to preserving actual works that UK cultural heritage institutions are looking to acquire. This will be done through small group exercises and will draw on approaches based on work carried out by Tate, the UK Legal Deposit Libraries (LDLs), and the Digital Preservation Coalition (DPC).

Keywords – file formats, time-based media, technology watch, collaboration, capacity building

Conference Topics – The Cutting Edge: Technical Infrastructure and Implementation; Exploring New Horizons.

I. LEARNING GOALS

- Participants will collaborate and exchange knowledge and experiences with other group members to enhance community understanding of this problem and to further build a definition of 'complex digital objects'.

- Participants will gain practical know-how to get started in planning for the preservation of complex digital objects at their home institutions.

II. DESCRIPTION

At the heart of preserving digital objects, particularly of non-standard or experimental content types, lies an inherent paradox that cannot be resolved by any single sequence of preservation actions, reference model, tool, or service. This paradox arises from the new and unprecedented types of content that can be generated by creators using innovative new technology.

These new and unprecedented works can be considered 'complex digital objects', which are:

- Born-digital with no print counterpart. They are defined by their native format and their intended access environment

- Constituted by formats that are complex and often networked and made of components that often consist of more than one media type

- Comprised of non-standard format and metadata types that might never become standardized

- Device-dependent and often require proprietary devices and platforms to enable the intended delivery of content and a meaningful user experience

- Not typically part of existing collections. Even mature collecting institutions might not have identified the necessary capability and infrastructure to manage them

- At risk of rapid obsolescence due to the transient and rapid change of the digital marketplace

While the resulting innovative objects provide novel ways for a creator to realize their vision, engage in the digital marketplace, and reach wider audiences, they also pose significant challenges for the institutions and individuals tasked with ensuring their preservation and access. No matter how up-to-date, responsive, and well-resourced an institution's response to digital preservation might be, the knowledge needed to

manage and preserve these objects will always lag behind the growth of the technology used in their creation.

This workshop addresses this paradox head-on. The organizers will apply the research they have undertaken in this area to small group activities. This approach will help to engage members of the digital preservation community to cultivate shared knowledge and to anticipate similar challenges that their institutions will encounter. While the 2018 iPres workshop 'Preservation of a Collaborative Community-Based Virtual Reality Collection' focused on strategies for engaging with the community to tackle the challenges of virtual reality objects, this workshop will more broadly address common challenges to preserving an array of different types of complex digital objects.

III. BACKGROUND

A. UK LDLs' Emerging Formats project

The UK LDLs' Emerging Formats project was set up to investigate the digital publishing landscape in the UK for more complex works that are in scope to acquire but not currently collected under the UK's Non-Print Legal Deposit Regulations. Within the scope of the project, the LDLs decided to focus on three formats: eBooks created as mobile apps, web-based interactive narratives, and structured data. The project used a sample of publications to determine how the libraries could acquire, preserve, and provide access to works created in these formats. Based on project findings, the LDLs created a methodology to inform how to address complex publications. This research is informed by existing digital preservation practices as well as new approaches that have come out of the project.

B. Tate's Time-based Media Conservation

Tate's Time-based Media (TiBM) conservation team is responsible for the preservation of Collection artworks using performance, film, slides, video, audio, and software. In some cases, the object of preservation is not necessarily the software or data but the experience of the artwork. To address this issue, the TiBM team has developed risk assessment and analysis processes to evaluate the vulnerability of individual artworks. The team weighs up the diverse options for preservation (from storage to migration and emulation). They then take steps to document the artwork and its technical history, while also making the work more sustainable, pre-empt future issues, and/or intervene to maintain the work's functions in the present.

These strategies at the UK LDLs and Tate provide a model, and in some cases tools, documentation, and procedures - that can be adapted (or even re-used outright) by other institutions.

IV. CONTENT

This workshop will discuss definitions for complex digital objects and provide an overview of the known challenges to preserving them. The first section of the workshop will focus on three predominant challenges:

1) Defining the digital object and its significant properties and using this information to decide what to preserve.

2) Problem-solving technical dependencies, including software and hardware environments.

3) Strategizing for digital rights management and intellectual property rights.

The organizers will present two to three case studies that exemplify these challenges.

Participants will then break out into small groups for activities designed to analyze and problem-solve the challenges of preserving complex digital objects. They will be asked to indicate their preferred activity from a choice of four different options while arriving at the workshop so that groups can be pre-arranged during presentations. The activity options will include:

1) an advocacy exercise to create a press release directed at building a preservation program;

2) a digital preservation workflow planning exercise;

3) a risk management exercise aimed at evaluating vulnerabilities; and

4) a donor agreement exercise that addresses significant properties, rights, and licensing issues.

The small group activities will be directed by Sara Day Thomson who co-delivers digital preservation training at multiple skill levels and across multiple professional sectors. Activity structure and worksheets will be adapted from the DPC's training resources.

In the final 30 minutes of the workshop, participants will feed back the results of their small group activities and discuss common trends as well as divergent approaches. Feedback will be collected and recorded in order to document the ideas and analysis generated by participants. This feedback will be shared with participants and published in the conference proceedings. The workshop will aim to identify opportunities for collaboration in the development of new approaches.

iPRES 2019 - 16th International Conference on Digital Preservation
September 16- 20, 2019, Amsterdam, The Netherlands.

2