INTERNATIONAL JOURNAL OF DIGITAL CURATION

Selecting Efficient and Reliable Preservation Strategies

Micah Altman
MIT

Richard Landau
MIT

DOI: https://doi.org/10.2218/ijdc.v15i1.727

ABSTRACT

This article addresses the problem of formulating efficient and reliable operational preservation policies that ensure bit-level information integrity over long periods, and in the presence of a diverse range of real-world technical, legal, organizational, and economic threats. We develop a systematic, quantitative prediction framework that combines formal modeling, discrete-event-based simulation, hierarchical modeling, and then use empirically calibrated sensitivity analysis to identify effective strategies.

Specifically, the framework formally defines an objective function for preservation that maps a set of preservation policies and a risk profile to a set of preservation costs, and an expected collection loss distribution. In this framework, a curator's objective is to select optimal policies that minimize expected loss subject to budget constraints. To estimate preservation loss under different policy conditions optimal policies, we develop a statistical hierarchical risk model that includes four sources of risk: the storage hardware; the physical environment; the curating institution; and the global environment. We then employ a general discrete event-based simulation framework to evaluate the expected loss and the cost of employing varying preservation strategies under specific parameterization of risks.

The framework offers flexibility for the modeling of a wide range of preservation policies and threats. Since this framework is open source and easily deployed in a cloud computing environment, it can be used to produce analysis based on independent estimates of scenario-specific costs, reliability, and risks.

We present results summarizing hundreds of thousands of simulations using this framework. This exploratory analysis points to a number of robust and broadly applicable preservation strategies, provides novel insights into specific preservation tactics, and provides evidence that challenges received wisdom.

PDF

PUBLISHED

29-Sep-2020

ISSUE

Vol 15 No 1 (2020)

SECTION

Conference Pre-prints

Open Journal Systems

INFORMATION

For Readers

For Authors

For Librarians

CURRENT ISSUE

ATOM 1.0

RSS 2.0

RSS 1.0