

READING THE MATRIX

A Hackathon Linking Digital Forensics To User Access

Seth Anderson

Yale University Library

USA

seth.r.anderson@yale.edu

Alex Chassanoff

Educopia Institute

USA

alex@educopia.org

Ethan Gates

Yale University Library

USA

ethan.gates@yale.edu

Abstract - A proposal for a multi-day hackathon event on the topic of bridging forensic disk imaging and bit-for-bit digital preservation workflows to emulation-based access services. Participants from a variety of backgrounds can collaborate to develop resources that ease, integrate and forefront access in the practice of digital forensics.

Keywords - Digital forensics, disk imaging, access, emulation

Conference Topics - The Cutting Edge: Technical Infrastructure and Implementation

I. INTRODUCTION

A. *Statement of Intent*

Extensive professional work and consideration has gone into using and documenting forensic tools for technical, bit-for-bit preservation of born-digital materials. Less discussed are practical solutions for how, once legacy material has been made physically and intellectually stable, to make it accessible. While emulation can be a highly successful strategy, it is frequently not obvious or easy to incorporate it into existing archival access services. How do we return raw bytes back to users for meaningful interaction?

We therefore propose a hackathon event on the topic of explicitly encouraging and linking forensic disk imaging and description to user access via emulation. Participants will be invited to work intensively during the

hackathon to deliver (or expand upon a pre-existing) a tool, document, diagram, or other resource that bridges this gap.

Potential projects might include: scripts to convert forensic disk images into a format mountable in emulation software; contributions to an existing open source emulation project (QEMU, SheepShaver, VICE); a tool to extract operating system metadata out of a disk image.

The iPres community is an ideal group to both inform and benefit from this hackathon, as the international digital preservation community's diverse expertise and experience with existing infrastructure and workflows is critical to identifying where current limitations lie.

B. *Planning Team*

This hackathon will be planned and organized by staff from the Educopia Institute and Yale University Library's Digital Preservation Services department, including the authors. Others will be available for on-site facilitation pending attendance.

Educopia and YUL, as key affiliates and sponsors of projects such as the BitCurator Consortium, the Software Preservation Network, and Scaling Emulation and Software Preservation Infrastructure (EaaS), are not just experts in the relevant technical practice of digital preservation described above, but in

community organization and creating collaborative learning environments. The organizers are well positioned to bring together and productively engage a diverse set of iPres members, attendees, and interested onlookers, in terms of both on-site logistics and leveraging any and all resources generated during the hackathon.

II. STRUCTURE

A. *Call for Participation*

The iPres community will be invited a month before the start of iPres 2019 to consider the topic as outlined above and submit their intent to participate, along with their potential hackathon project ideas via Google Form. The responses for project ideas specifically will be made publicly visible, to encourage brainstorming and allow participants to build off of each other's ideas in advance.

The Call for Participation will emphasize that attendance will be free for all and available to remote participants.

Filling out the participation form will also allow organizers to send an invitation to a dedicated Slack workspace, which will be used by organizers and participants to communicate and brainstorm with each other before and during the hackathon.

B. *Hack Day*

The central focus of the hackathon will be a day-long (approx. 8-hour) intensive "Hack Day", in which participants divide themselves into small groups (ideally 3-8 individuals each)

During the first half hour of the Hack Day, organizers will review previously submitted project proposals, allow for attendees to submit new proposals on the spot, and facilitate attendees splitting into groups and starting work based on their interests.

Organizers will periodically check in with groups throughout the day to encourage progress toward a deliverable resource or document, check on the logistics of remote participation, and otherwise field requests.

A stable wifi network will be necessary.

C. *Conference Follow Up*

Following the intensive Hack Day and for the duration of iPres 2019, until an allotted wrap-up presentation slot, hackathon groups will be free to communicate with each other and continue working on or refining their deliverable as desired.

Organizers will confirm that each group has at least one designated representative to present the group's work to the wider iPres community at a closing session presentation. Organizers will facilitate the session and, following the conference, confirm with each group that their deliverable resources are uploaded and accessible on a public platform.

The wrap-up presentation will include an "audience award" segment allowing attendees to vote on and reward outstanding projects. Prizes for winning groups will be arranged.

D. *Remote Participation*

Using Slack will enable remote hackathon participants to join and communicate with groups. At least one hackathon organizer will be a remote participant, to ensure organizers facilitate remote inclusion in group projects.

III. GOALS

A. *Create Resources*

By the end of the conference, the hackathon will produce a set of concrete resources, with each group determining their exact deliverables (a script or code, workflow documentation, data set, etc.)

Each group's deliverable resources(s) must be open sourced and made available to the iPres community.

B. *Engage and Bridge Communities*

The hackathon will provide a multi-day space for specialists in related but too-often disparate sides of digital preservation – preservation and access services – to connect one-on-one and productively tackle obstacles to user access of legacy digital materials. The conversation and networking involved will spark further, in-depth collaborations.