2021-08-17 Whitman College Migration Pilot Check-in

Attendees

- Alan Stanley
- Amy Blau
- Dana Bronson
- ~~Dara Virks~~
- David Wilcox
- Kun Lin
- Paige Morfitt

Agenda

Topic
Stand-up 1. What has everyone been working on? 2. What will you be working on over the next couple weeks? 3. Are there any blockers that are preventing you from getting work done?
Production Migration Schedule
Priorities and areas of focus 1. Timeline for remaining work 2. Possible cut-over date 3. Documentation tasks
Wrap-up and next steps

Notes

1. Alan
    a. Wrote a Python script to parse the logs
        i. Getting lots of 500 errors on the server, the script finds these and fixes the errors
    b. Can't delete media because we can't delete files
        i. Alan thinks this is a setting we can add and then remediate afterward
        ii. Amy wants to make sure Whitman team can clean these up on their own afterward
    c. Issues with generating thumbnails from large PDFs
        i. May have to generate by hand for now
    d. Space issues
        i. Files do not seem to be duplicated in EBS and Fedora anymore
2. Amy
    a. Testing with limited access items
        i. Spreadsheet with some of the theses
        ii. Would like to try on staging first
        iii. Can use HTTP links from an S3 bucket with IP restrictions in place
    b. Pace of migration
        i. Alan is down to the newspapers
        ii. Each newspaper spreadsheet is taking about a day
            1. Alan put a 6 second pause in between pages
            2. If the OCR doesn't transfer the first time it tries again and gets it the second time
            3. Only about 50 errors so far
    c. Cut-over
        i. Still need some documentation and training
            1. Go through documentation and identify specific questions
        ii. CAS sign-in
            1. Need to make this a little more user-friendly
        iii. Need to verify limited access items
        iv. Hopefully within the next 2 weeks

Actions