Greater Than the Sum?

Challenges of Integrating Multiple Digital Media Types Into a Unified Preservation and Access System

March 11, 2016
ARLIS/NA+VRA 3rd Joint Conference, Seattle, WA

Geoff Froh, Deputy Director
(geoff.froh@densho.org)
Greater Than the **Checksum**?

A Different Approach to the Challenge of Storing and Preserving Complex Media Collections

March 11, 2016
ARLIS/NA+VRA 3rd Joint Conference, Seattle, WA

Geoff Froh, Deputy Director
(geoff.froh@densho.org)
“Issei machinists. c.1918.” ddr-densho-12-17
Courtesy of Densho and the Kunitsugu Family Collection.
Densho: to pass on to the next generation or leave a legacy

Visual Histories

956 interviews
871 narrators
1880 hours video/audio

All 10 camps; 14 “assembly centers”
26 birth states; 5 countries
Photos & Documents

~168,000 files
345 collections
7 partners

Photographs, documents, periodicals, ephemera
High-quality mezzanine files
Creative Commons (CC-NC-SA) license
Introducing git

- Distributed source control system for developers
- “Repositories” (“repos”) are just directories
- Change-tracking and logging built in
- Repos can be cloned from and synced with each other

More info - https://git-scm.com/
...and git-annex

- Allows large binaries to be stored in git repo
- Automatically checksums files
- Clones do not have to contain all annex files
- Tracks where files are located
- Supports local and cloud storage backends

More info - https://git-annex.branchable.com/
JSON file format

- Structured key/value text files
- Widely supported by software; but human-readable

```json
{
    "Person": {
        "last_name": "Parkins"
    },
    "first_name": "Fred"
}
```
Simplifying metadata with JSON

- Collection-level: EAD/DACS-ish
- Object-level: METS/MODS/PREMIS-ish

```json
{
    "id": "ddr-densho-23-1"
},
{
    "title": "Children from Nihonmachi"
},
{
}
```
Putting it together: A new model

```
ddr-densho-8
 |__ collection.json
    |__ objects
    |__ ddr-densho-8-1
       |__ object.json
       |__ files
           |__ ddr-densho-8-1-78df31-transcript.pdf
           |__ ddr-densho-8-1-b988da-master.mpg
           |__ ddr-densho-8-1-9003fb-mezz.mp4
           |__ ddr-densho-8-1-2246ac-screen.jpg
    |__ ddr-densho-8-2
       |__ object.json
...
```

Collection = repo
DDR
Moving data
DDR

Publishing a collection
DDR Platform

Logical model

"Workbench"
ddr-editor

"Core"
ddr-core

"Public"
ddr-public
“Masao Sakagami looking out at a view of mountains.” ddr-densho-201-470
Courtesy of Densho and the Masao Sakagami Collection.
Thank you!

www.densho.org
ddr.densho.org
archive.densho.org

Questions? Followup?
Geoff Froh (geoff.froh@densho.org)