

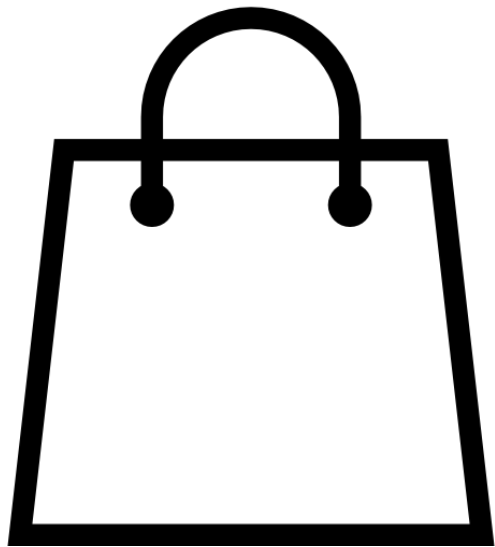
HOW + WHEN TO BAGIT

NDSR-NYC Symposium | April 28, 2016

KATHRYN GRONSBELL + GENEVIEVE HAVEMEYER-KING

bit.ly/1qQMVue

- Activity instructions
- Baglt Resources



**WHAT IS
BAGIT?**

WHAT IS BAGIT?

- Hierarchical package format
- For storage and transfer
- Specification by:
 - Library of Congress
 - California Digital Library
- FREE
- Current: V0.97, 01/26/2016

* <https://tools.ietf.org/html/draft-kunze-bagit-13>

WHAT DOES BAGIT **DO**?

- By itself? **Nothing!**
- Establishes a structure
- Describes itself
- Enables validation

VALIDATION means...

FIXITY + **ATTENDANCE**

= data you have has not
changed in any way






(bit level)

= files exist exactly as they
were captured

(not deleted, moved, added)

BAGIT EXAMPLE

payload

Name	
	bag-info.txt
	bagit.txt
	data
	manifest-md5.txt
	tagmanifest-md5.txt

WHY NOT **D.I.Y.** A BAG?



WHY USE BAGIT?

- Digital preservation
- Flexible spec = **holey bags**
 - URLs stored in payload, not data
- Used in workflows at:
 - MoMA
 - AFC
 - Brooklyn Academy of Music
 - MetaArchive / Archivematica users...

BIG BAGS

- **Scale & Documentation**
- **Many Bytes vs. Many Files**
- **Workstation capabilities?**
- **Options for Large Bags:**
 - **Holey bags?**
 - **Split Bags?**

BREAK

bit.ly/1qQMVue

**ALL
GOOD!**

**HELP
ME!**

BAGIT LIBRARIES AND IMPLEMENTATIONS

JAVA:

- Most common implementation (e.g. in Archivematica)
- Newest release (5) does **NOT offer CLI support**

PYTHON:

- Offers CLI options
- Enables programmable usage (aka “library usage”)
 - scripts to create/modify bags and bag metadata

**WE ARE WORKING ON
THE COMMAND LINE
AND...**

WE ARE MODIFYING DATA

create-bag.md

Creating a bag means:

- packaging and describing/creating manifests for data
- structuring data

THIS WILL REPLACE YOUR DATA a.k.a. “BAG IN PLACE”

validate-bag.md

Validation means:

- Bag is **COMPLETE**
 - Required elements present + structured to spec
 - Payload data + txt files appear in manifest
- **AND** all checksums for all files in bag can be validated

more-bag-activities.md

- **Choose which additional activities interest you**
 - **Or just use this time to create/validate more bags!**
- **Link to BagIt resources (please add more!)**

**What was
most
helpful?**

**How can
we
improve?**

bit.ly/1qQMVue

THANK YOU!

[@k_grons](#)

[@genevieve_hk](#)