Electronic Records Processing: It's a CINCH!

Dean Farrell, Systems Integration Librarian, State Library of North Carolina

What is it?

CINCH (Capture INgest CHecksum) is a tool that automates the transfer of online content to a repository, using ingest technologies appropriate for digital preservation.

More familiarly, CINCH

- grabs freely available online content,
- authenticates it,
- extracts metadata, and
- readies it for repository ingest.

Why?

The State Library of North Carolina has a legislative mandate to collect and preserve print and digital state government publications.

Current workflow for born-digital publications?

- download files one-by-one
- receive via email or CD
- process files with multiple manual steps

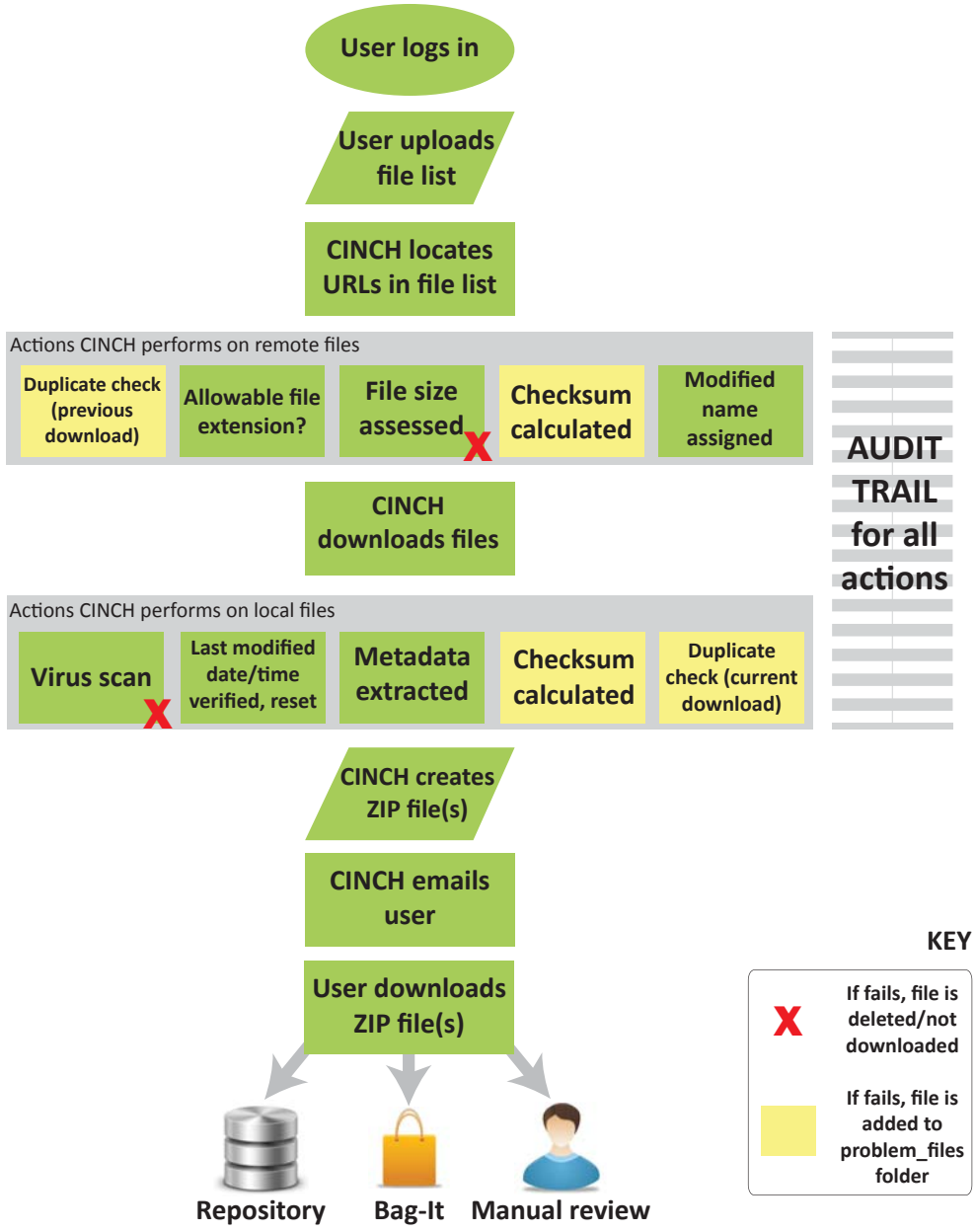
Now, CINCH:

- automates batch file downloads from the web &
- follows digital preservation best practices along the way.

Institute of Museum and Library Services

North Carolina Department of Cultural Resources

How does it work?



CINCH is...

- free
- repository-neutral
- lightweight
- easy to use
- open source

Supported file types?

.csv	.png
.doc & .docx	.ppt & .pptx
.gif	.txt
.pdf	.xls & .xlsx

What's behind CINCH?

- PHP
- MySQL
- Yii Framework
- Apache Tika – metadata extraction
- ClamAV – virus checking

Where can I get it?

http://slnc-dimp.github.com/Cinch

North Carolina institutions will be able to use a hosted version in August, 2012.

Funding for the CINCH: Capture, INgest, & CHecksum tool is made possible through an IMLS Sparks! Ignition grant.