Introduction

• Clowder is a customizable and scalable data management system you can install anywhere.
• You can install Clowder in the cloud, on your hardware, or you can partner with NCSA for a custom instance.
• You can contribute to the core software or by creating new metadata extractors and data visualizations.

Flexible Metadata Representation

Support for both user-defined and machine-defined metadata. System accepts metadata in a flexible representation based on JSON Linked Data.

Users can add metadata entries directly from the user interface.

Extractors and external clients can attach metadata to files and datasets using the Web service API.

Data Visualizations

• To preview the content of large files and visualize the information contained in files and datasets in a meaningful way, Clowder provides ways to write Javascript-based visualizations.
• Often these data previews are added by automatic extractions.

Organize Data

Spaces (top-left), Datasets (top-right), Files (bottom-left), Collections (bottom-right).

Automatic Metadata Extraction

Extend the system by creating new extractors to analyze data. For example, given the image on the right 3 separate extractors computed:
• EXIF from image bytes
• Siegfried file format identification
• OCR of text in image

Write your extractors using the Clowder Python SDK or in most other programming languages.

Search

Flexible text search over basic attributes of files, datasets and collections. Advanced search over metadata fields.

Open Source Community

Used across many different research areas.

Clowder has been funded by: NSF #BCS-0941268, #EAR-331906, #ACI-1261582, #ACI-1443013, #ACI-0940824, #ACI-0525308, #OAC- 1835834; NIH #1P01AI089556-01A1. Illinois - Indiana SeaGrant #DW92329201, European Commission #PR-261600, XSEDE #ACI-1053575, ARPA-E #DE-AR0000594.