

# Partial Extractor List

This will be replaced by <https://catalog.clowderframework.org/>

## Scientific

Domain	Tool	Description	Repositories
Hydrology	<b>Advection Diffusion</b>	Solve a general advection-dispersion equation.	
	<b>Chemical Mean Age</b>	Determine the mean age of chemical constituents with inputs of chemical dynamics.	
	<b>Document Tables Extractor</b> (nca.nlp.wordtables)	Extract tables from documents.	<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/WordTablesExtractor">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/WordTablesExtractor</a>
	<b>GDAL</b> (nca.geo.shpExtractor)		<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-geo/browse">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-geo/browse</a>
	<b>GDAL</b> (nca.geo.tiffExtractor)		<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-geo/browse">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-geo/browse</a>
	<b>GDAL</b> (nca.image.geotiff)		<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-geotiff/browse">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-geotiff/browse</a>
	<b>Historical River Extractor</b> (nca.cv.river)	Extract the river networks from the ancient hand-drawing maps and compare them with current river networks	<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/river">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/river</a>
	<b>Normalized Difference Vegetation Index</b> (nca.arctis.landsat7mosaic)		<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-bd-cz/browse/ndviextractor">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-bd-cz/browse/ndviextractor</a>
	<b>River Chi Index</b>	Identify the river dynamics in a river basin and evaluate human activities' influences through Chi index in the streams.	<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-bd-cz/browse/chi-analysis">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-bd-cz/browse/chi-analysis</a>
	<b>River Sinuosity</b>	Study the maturity and equilibrium conditions of a stream through the sinuosity index.	
	<b>Soil Moisture Change</b>	Determine role of hydraulic redistribution in AZ (riparian site / upland site) by studying soil moisture change throughout different seasons.	
	<b>Species Classifier</b>	SAM based Species Classification from <b>Hyperspectral</b> data, Hyperspectral Indices, NDVI, SAVI, MSAVI, etc.	
	<b>TerEx</b> (nca.arctis.floodplain)	Identify the flat polygons and the heights inside a river valley.	<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-bd-cz/browse/terex_floodplain">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-bd-cz/browse/terex_floodplain</a>
	<b>Topographic Depressions</b>	Identify topographic depressions (TDs) and their distribution on landscape (Number, location, area, volume of TDs).	<a href="https://github.com/HydroComplexity/TDI">https://github.com/HydroComplexity/TDI</a>
<b>Valley Safety Zones</b>	Estimate submerging areas and water depths under extreme floods and map the safety zones in a river valley.		
<b>Vegetation Indices</b>	Calculating vegetation indices such as NDVI and Surface Temperature from Landsat 7 and 8 <b>satellite</b> data.		
Ecology	<b>PlantCV</b> (terra.plantcv)	Extract plant height, area, and color distribution from photographs.	<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-plantcv/browse">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-plantcv/browse</a>
Civil & Environmental Engineering	<b>Body of Water Detector</b> (nca.image.ponddetect)	Land coverage, extract locations of bodies of water from <b>satellite</b> data.	<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-maps/browse/feature_detection">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-maps/browse/feature_detection</a>
	<b>GI Identification</b>		
	<b>Human Preference Score</b> (nca.image.humanpref)	Assign a model derived human preference score to a given image of an urban environment.	<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-maps/browse/humanpref">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-maps/browse/humanpref</a>
	<b>Route Greenness</b> (nca.xml.greenindexroute)	Derive the green index of a city route.	<a href="https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-maps/browse/greenroute">https://opensource.nca.illinois.edu/bitbucket/projects/CATS/repos/extractors-maps/browse/greenroute</a>
	<b>Social Media GI Preferences</b>		

	<b>Stanford CoreNLP - Sentiment</b> (nca.nlp.SNLPSentiment)	Assign a sentiment score to a piece of text.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/SNLP/SNLPSentimentExtractor">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/SNLP/SNLPSentimentExtractor</a>
	<b>TUV Triaxus</b>		
<b>Social Science &amp; Humanities</b>	<b>Bertillon Card Cell Extractor</b> (nca.image.dmp)	Extract table cells from a Bertillon Card.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/DEBOD/repos/extractors-debod">https://opensource.ncsa.illinois.edu/bitbucket/projects/DEBOD/repos/extractors-debod</a>
	<b>Census Form Cell Extractor</b> (census-section-segmentor)	Extract table cells from a 1930s Census form.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/census">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/census</a>
	<b>Handwritten Decimals Extractor</b> (nca.image.sphog.debod)	Extract handwritten decimal values from an image.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/DEBOD/repos/extractors-handwrittendecimals">https://opensource.ncsa.illinois.edu/bitbucket/projects/DEBOD/repos/extractors-handwrittendecimals</a>
	<b>Killed Photos</b> (nca.image.iarp_remove_circle)	Identify depression era photos "killed" by Farm Security Administration director Roy Stryker, indicated by a hole punch in the image.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/IARP/repos/image_fetcher/browse/extractors/remove_circle">https://opensource.ncsa.illinois.edu/bitbucket/projects/IARP/repos/image_fetcher/browse/extractors/remove_circle</a>
	<b>Mean Grey</b> (nca.cv.meangrey)	Mean grey values of black and white photos.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/IARP/repos/image_fetcher/browse/extractors/mean_grey">https://opensource.ncsa.illinois.edu/bitbucket/projects/IARP/repos/image_fetcher/browse/extractors/mean_grey</a>
	<b>Movie Slice</b> (nca.movieslice)		<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-movieslice/browse">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-movieslice/browse</a>
	<b>Person Detector</b> (person-detector)	Extract locations of people in an image.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-person-detector">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-person-detector</a>
	<b>Person Tracker</b> (nca.person-tracker)	Extract locations and paths of people moving in videos.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-person-tracking">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-person-tracking</a>
	<b>Video Analytics Toolbox</b> (nca.cinematics_batch)	Extract shot descriptors from videos.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-lsva-integrated/browse">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-lsva-integrated/browse</a>
<b>Biology, Genomics, Medicine</b>	<b>FSL</b> (mri2mesh)		<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-mri/browse/mri2mesh">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-mri/browse/mri2mesh</a>
	<b>Glomeruli</b> (nca.msc.diagnosis)	Extract glomeruli from Kidney biopsy images.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-msc/browse/Kidney">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-msc/browse/Kidney</a>

## General

Domain	Tool	Type	Description	Repositories
	<b>CMU Sphinx</b> (nca.audio.speech2text)	Extract or	Audio recognition to extract text for speech within audio.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-core/browse/audio/speech2text">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-core/browse/audio/speech2text</a>
	<b>DBpedia</b> (nca.dbpedia)	Extract or	Find and define named entities within the given text.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-dbpedi/browse">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-dbpedi/browse</a>
	<b>ImageMagick</b> (nca.image.metadata)	Extract or	Pull available EXIF image metadata from a given image.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-core/browse/image/metadata">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-core/browse/image/metadata</a>
	<b>OpenCV - Faces</b> (nca.cv.faces)	Extract or	Find faces in an image and return their locations.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-opencv/extractors-opencv-faces">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-opencv/extractors-opencv-faces</a>
	<b>OpenCV - Eyes</b> (nca.cv.eyes)	Extract or	Find eyes in an image and return their locations.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-opencv/extractors-opencv-eyes">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-opencv/extractors-opencv-eyes</a>
	<b>OpenCV - Closeups</b> (nca.cv.closeups)	Extract or	Determine whether an image is a closeup of a person or not.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-opencv/extractors-opencv-closeups">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-opencv/extractors-opencv-closeups</a>
	<b>OpenCV - Profiles</b> (nca.cv.profiles)	Extract or	Find human face profiles in an image and return their locations.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-opencv/extractors-opencv-profiles">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-opencv/extractors-opencv-profiles</a>
	<b>Langid</b> (nca.nlp.simplelanguage)	Extract or	Identify the language of the given text.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/SimpleLanguage">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/SimpleLanguage</a>
	<b>NLTK - Summary</b> (nca.nlp.simplesummary)	Extract or	Summarize a body of text.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/SimpleSummary">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/SimpleSummary</a>
	<b>Siegfried</b> (siegfried)	Extract or	Extract information about a given file relevant to identifying its type and validating its format.	<a href="https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-siegfried/browse">https://opensource.ncsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-siegfried/browse</a>

<b>General</b>	<b>Stanford CoreNLP</b> (nlsa.nlp.SNLP)	Extract or	Natural Language Process extractions such as parts of speech, named entities, language, etc.	<a href="https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/SNLP/SNLPExtractor">https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/SNLP/SNLPExtractor</a>
	<b>Tesseract</b> (nlsa.image.ocr)	Extract or	Object Character Recognition (OCR) to extract text from images containing text.	<a href="https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-tesseract">https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-tesseract</a>
	<b>Tika</b> (nlsa.nlp.tika)	Extract or	Document extractions such as language identification, ...	<a href="https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/tika">https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-nlp/browse/tika</a>
	<b>Versus - Color Distribution</b> (nlsa.versus.image)	Extract or	Generate a distribution of color values within an image to be used for comparing how similar two images are.	<a href="https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-versus/browse">https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-versus/browse</a>
	<b>VLFeat</b> (nlsa.image.caltech101)	Extract or	Classify images as to whether they contain objects from the Caltech101 dataset (e.g. people, airplanes, motorcycles, cougars, ...).	<a href="https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-vlfeat">https://opensource.nsa.illinois.edu/bitbucket/projects/CATS/repos/extractors-cv/browse/extractors-vlfeat</a>

Older *Extractors*