

# Project Supported Transformations Progress

**Code:** Is there code? Is it submitted or existing? If so add URL to parent wiki page!!!

**Py1/Py2:** Is it using PyCloudwer1, PyCloudwer 2, or neither. Expectation for Beta is most will be in Py1, moving forward with new ones we will use Py2.

**JSON-LD:** Is proper JSON-LD generated in the MetaData and registration document (make sure to use full URI).

**DCKR:** Dockerized?

**Test File:** Sample file available?

**TC:** Is the tool in the tools catalog?

**DPL:** Is the tool dockerized (if not Windows), managed by the elasticity service, and passing tests every other hour?

**DL:** Is the tool downloadable via Sandeep's modifications to move compute to the data?

|                     |
|---------------------|
| Complete & Deployed |
| Needs work          |
| Blocked             |

## Scientific

| Domain    |    | Tool                                                                                      | Type      | Description                                                                                                                                     | Code                                                                                        | Py1/Py2 | JSON-LD | DCKR                         | Test File | TC  | DPL                     | DL  | Assignment                                                            |
|-----------|----|-------------------------------------------------------------------------------------------|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|---------|---------|------------------------------|-----------|-----|-------------------------|-----|-----------------------------------------------------------------------|
| Hydrology | 1  | Advection Diffusion                                                                       |           | Solve a general advection-dispersion equation.                                                                                                  | No<br><br>(Don't have a general tool. Seems like this was not done as per the description.) |         |         |                              |           |     |                         |     | Sandeep Puthanveetil Satheesan<br><br>Per Dr K this will be post beta |
|           | 2  | Chemical Mean Age                                                                         | Extractor | Determine the mean age of chemical constituents with inputs of chemical dynamics.                                                               | No<br><br>(This is a simulation tool - don't have code.)                                    |         |         |                              |           |     |                         |     | Sandeep Puthanveetil Satheesan<br><br>Per Dr K this will be post beta |
|           | 3  | Document Tables Extractor (nrsa.nlp.wordtables)                                           | Extractor | Extract tables from documents.                                                                                                                  | Yes                                                                                         | No      | No      | No                           | No        | No  | No                      | No  |                                                                       |
|           | 4  | ESRI Shapefile Extractor by using GDAL (nrsa.geoshp.preview)                              | Extractor | Create a Geoserver layer for WMS service. It can be used in previewer                                                                           | Yes                                                                                         | Py1     | Yes     | Yes                          | Yes       | No  | Yes                     | No  | Jong Lee                                                              |
|           | 5  | GeoTiff Extractor by using GDAL (nrsa.geotiff.preview)                                    | Extractor | Create a Geoserver layer for WMS service. It can be used in previewer                                                                           | Yes                                                                                         | Py1     | Yes     | Yes                          | Yes       | No  | Yes                     | No  | Jong Lee                                                              |
|           | 6  | GeoTiff Metadata Extractor by using GDAL (Another implementation) (nrsa.geotiff.metadata) | Extractor | Extract geospatial metadata from Geotiff. It should be combined with #5 Geotiff extractor.                                                      | Yes                                                                                         | Py1     | Yes     | Yes                          | Yes       | Yes | Yes                     | No  | Jong Lee                                                              |
|           | 7  | Historical River Extractor (nrsa.cv.river)                                                | Extractor | Extract the river networks from the ancient hand-drawing maps and compare them with current river networks                                      | Yes                                                                                         | Py1     | Yes     | Yes                          | Yes       | Yes | Yes                     | No  | Sandeep Puthanveetil Satheesan                                        |
|           | 8  | Normalized Difference Vegetation Index (nrsa.arccis.landsat7mosaic)                       | Extractor | Create a NDVI layer from Landsat data. Calculating vegetation indices such as NDVI and Surface Temperature from Landsat 7 and 8 satellite data. | Yes                                                                                         | Py1     | Yes     | N/A<br><br>(Windows, ArcGIS) | Yes       | No  | Yes<br><br>(Windows VM) | No  | Sandeep Puthanveetil Satheesan                                        |
|           | 9  | River Chi Index (nrsa.chi_analysis)                                                       | Extractor | Identify the river dynamics in a river basin and evaluate human activities' influences through Chi index in the streams.                        | Yes                                                                                         | Py1     | Yes     | Yes                          | Yes       | Yes | Yes                     | Yes | Sandeep Puthanveetil Satheesan                                        |
|           | 10 | GDAL Converter                                                                            | Converter | Converts between TIFF and BIL (zip) and FLT (zip)                                                                                               | Yes                                                                                         | N/A     | N/A     | Yes                          | Yes       | Yes | Yes                     | Yes | Sandeep Puthanveetil Satheesan                                        |

|         |    |                                                              |           |                                                                                                                                                                                                                               |                                         |     |     |                       |     |     |                  |     |                                                                                                                   |
|---------|----|--------------------------------------------------------------|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------|-----|-----|-----------------------|-----|-----|------------------|-----|-------------------------------------------------------------------------------------------------------------------|
|         | 11 | River Sinuosity                                              | Extractor | Study the maturity and equilibrium conditions of a stream through the sinuosity index.                                                                                                                                        | Yes                                     |     |     |                       |     |     |                  |     | Sandeep Puthanveetil Satheesan<br><br>Qina supposed to be delivering code - Dr K thinks it should make it in Beta |
|         | 12 | Soil Moisture Change                                         | Extractor | Determine role of hydraulic redistribution in AZ (riparian site / upland site) by studying soil moisture change throughout different seasons.                                                                                 | No                                      |     |     |                       |     |     |                  |     | Sandeep Puthanveetil Satheesan<br><br>Per Dr. K - will need discussed and can not be in Beta                      |
|         | 13 | Species Classifier                                           | Extractor | SAM based Species Classification from <b>Hy</b> perspectral data, Hyperspectral Indices, NDVI, SAVI, MSAVI, etc.                                                                                                              | Yes                                     |     |     |                       |     |     |                  |     | Sandeep Puthanveetil Satheesan<br><br>Dr K does not think this will make it in Beta                               |
|         | 14 | Floodplain Extractor by using TerEx (nasa.arcgis.floodplain) | Extractor | Identify the flat polygons and the heights inside a river valley.                                                                                                                                                             | Yes (but procedure?)                    | Py1 | Yes | No (windows)          | Yes | No  | Yes (Windows)    | no  | Bing Zhang                                                                                                        |
|         | 15 | Topographic Depressions                                      | Extractor | Identify topographic depressions (TDs) and their distribution on landscape (Number, location, area, volume of TDs).                                                                                                           | Yes - ready to review                   |     |     |                       |     |     |                  |     | Sandeep Puthanveetil Satheesan                                                                                    |
|         | 16 | Tree Delineation (nasa.arcgis.treedelin)                     | Extractor | Create a shape-file with polygons of tree canopy polygon from LIDAR data                                                                                                                                                      | Yes                                     | Py1 | Yes | N/A (Windows, ArcGIS) | Yes | Yes | Yes (Windows VM) | No  | Sandeep Puthanveetil Satheesan                                                                                    |
|         | 17 | Tree-wise voxelization                                       |           | Tree-wise voxelization of waveform data for <b>lidar</b> metrics that describes canopy structure (max intensity, height, etc...). Individual tree delineation, tree leaf area density to describe vertical leaf distribution. | Yes (But extractor has to be developed) | No  | No  | No                    | No  | No  | No               | No  | Sandeep Puthanveetil Satheesan                                                                                    |
|         | 18 | Valley Safety Zones                                          | Extractor | Estimate submerging areas and water depths under extreme floods and map the safety zones in a river valley.                                                                                                                   | Yes                                     |     |     |                       |     |     |                  |     | Sandeep Puthanveetil Satheesan<br><br>emailed March and April.<br><br>Sandeep will follow up with Qina.           |
| Ecology | 1  | netcdf (ncdump)                                              | Converter | Convert from binary netcdf to text.                                                                                                                                                                                           | N/A                                     | N/A | N/A | Yes                   | Yes | Yes | Yes              | Yes | Yan Zhao                                                                                                          |
|         | 2  | PEcAn (PEcAn#Ameriflux)                                      | Converter | Convert Ameriflux data to PEcAn's netcdf CF format.                                                                                                                                                                           | Yes                                     | N/A | N/A | Yes                   | Yes | Yes | Yes              | Yes | Yan Zhao                                                                                                          |
|         | 3  | PEcAn (PEcAn#DALEC)                                          | Converter | Convert PEcAn's netcdf CF format to the format required by the DALEC model.                                                                                                                                                   | Yes                                     | N/A | N/A | Yes                   | Yes | Yes | Yes              | Yes | Yan Zhao                                                                                                          |
|         | 4  | PEcAn (PEcAn#ED2)                                            | Converter | Convert PEcAn's netcdf CF format to the format required by the ED model.                                                                                                                                                      | Yes                                     | N/A | N/A | Yes                   | Yes | Yes | Yes              | Yes | Yan Zhao                                                                                                          |
|         | 5  | PEcAn (PEcAn#LINKAGES)                                       | Converter | Convert PEcAn's netcdf CF format to the format required by the LINKAGES model.                                                                                                                                                | Yes                                     | N/A | N/A | Yes                   | Yes | Yes | Yes              | Yes | Yan Zhao                                                                                                          |
|         | 6  | PEcAn (PEcAn#Sipnet)                                         | Converter | Convert PEcAn's netcdf CF format to the format required by the Sipnet model.                                                                                                                                                  | Yes                                     | N/A | N/A | Yes                   | Yes | Yes | Yes              | Yes | Yan Zhao                                                                                                          |

|                                   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |           |                                                                                         |                                                                                               |     |     |     |                                                                         |     |                                                    |    |                            |
|-----------------------------------|---|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|-----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|-----|-----|-----|-------------------------------------------------------------------------|-----|----------------------------------------------------|----|----------------------------|
|                                   | 7 | <p><a href="#">PEcAn</a><br/>(PEcAn#Ame<br/>rifluxBNL)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#FLU<br/>XNET2015)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#FAC<br/>E)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#PAL<br/>EON)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#NLD<br/>AS)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#CUR<br/>NCEP)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#GLD<br/>AS)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#GFD<br/>L)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#BIO<br/>CRO)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#CLM)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#GDA<br/>Y)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#JUL<br/>ES)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#LPJ-<br/>GUESS)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#MAA<br/>T)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#MAE<br/>SPA)</p> <p><a href="#">PEcAn</a><br/>(PEcAn#PRE<br/>LES)</p> |           |                                                                                         | Yes                                                                                           | N/A | N/A | No  | Yes                                                                     | Yes | Yes                                                | No | <a href="#">Yan Zhao</a>   |
|                                   | 8 | <a href="#">PlantCV</a><br>(terra.plantcv)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Extractor | Extract plant height, area, and color distribution from photographs.                    | Yes                                                                                           | Py1 | yes | yes | yes                                                                     | yes | no<br>(bamboo test available , not being deployed) | no | <a href="#">Yan Zhao</a>   |
|                                   | 9 | BIL (hyperspectral)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | Converter | Convert bil.zip (contains raw & *.hdr) to *.terra.nc                                    | Yes<br><br>(The original code has been updated. Need to figure out if we want to update this) | N/A | N/A | yes | no, only test files available are very large and would take a long time | yes | no<br>(bamboo test available , not being deployed) | no | <a href="#">Yan Zhao</a>   |
| Civil & Environmental Engineering | 1 | <b>Body of Water Detector</b><br>(ncaa.image.ponddetect)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | Extractor | Land coverage, extract locations of bodies of water from <b>satellite</b> data.         | Yes                                                                                           | No  | No  | No  | Yes                                                                     | No  | No                                                 | No | ??who??                    |
|                                   | 2 | <b>GI Identification</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | Extractor |                                                                                         | Yes                                                                                           | Py1 | Yes | Yes | Yes                                                                     | Yes | Yes                                                | No | <a href="#">Bing Zhang</a> |
|                                   | 3 | <b>Human Preference Score</b><br>(ncaa.image.humanpref)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Extractor | Assign a model derived human preference score to a given image of an urban environment. | Yes                                                                                           | no  | no  | no  | yes                                                                     | no  | yes                                                | no | ??who??                    |
|                                   | 4 | <b>Route Greenness</b><br>(ncaa.greenindex)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Extractor | Derive the green index of a city route.                                                 | Yes                                                                                           | Py1 | Yes | yes | yes                                                                     | Yes | Yes                                                | No | <a href="#">Bing Zhang</a> |
|                                   | 5 | <b>Social Media GI Preferences</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | Extractor | Determine if text contains references to visual or functional green infrastructure.     | Yes                                                                                           | Py1 | Yes | Yes | Yes                                                                     | No  | no                                                 | No | <a href="#">Bing Zhang</a> |

|                             |                                                            |                                                                   |                                       |                                                                                                                                       |                  |                    |     |                          |     |     |     |     |                                |                                |
|-----------------------------|------------------------------------------------------------|-------------------------------------------------------------------|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|------------------|--------------------|-----|--------------------------|-----|-----|-----|-----|--------------------------------|--------------------------------|
| Social Science & Humanities | 6                                                          | <b>Stanford CoreNLP - Sentiment</b><br>(nlsa.nlp.SNLPSentiment)   | Extractor                             | Assign a sentiment score to a piece of text.                                                                                          | Yes              | Py2<br>Java equiv. | Yes | Yes                      | Yes | Yes | No  | No  | Gregory Jansen                 |                                |
|                             | 7                                                          | <b>TUV Triaxus</b>                                                | Extractor                             | Towed Undulating Vehicle Data Analyzing Tools                                                                                         | Yes              | Py1                | yes | yes                      | yes | yes | yes | No  | Yan Zhao                       |                                |
|                             | 1                                                          | <b>Bertillon Card Cell Extractor</b><br>(nlsa.image.dmp)          | Extractor                             | Extract table cells from a Bertillon Card.                                                                                            | Yes              | Py1                | Yes | No                       | No  | No  | No  | No  | Sandeep Puthanveetil Satheesan |                                |
|                             | 2                                                          | <b>Census Form Cell Extractor</b><br>(census-section-segmentor)   | Extractor                             | Extract table cells from a 1930s Census form.                                                                                         | Yes              | No                 | No  | No                       | Yes | No  | No  | No  | Sandeep Puthanveetil Satheesan |                                |
|                             | 3                                                          | <b>Handwritten Decimals Extractor</b><br>(nlsa.image.sphog.debod) | Extractor                             | Extract handwritten decimal values from an image.                                                                                     | Yes              | Py1                | Yes | No                       | No  | No  | No  | No  | Sandeep Puthanveetil Satheesan |                                |
|                             | 4                                                          | <b>Killed Photos</b><br>(nlsa.image.killedphoto)                  | Extractor                             | Identify depression era photos "killed" by Farm Security Administration director Roy Stryker, indicated by a hole punch in the image. | Yes              | Py2                | Yes | Yes                      | Yes | Yes | no  | no  | Sandeep Puthanveetil Satheesan |                                |
|                             | 5                                                          | <b>Mean Grey</b><br>(nlsa.cv.meangrey)                            | Extractor                             | Mean grey values of black and white photos.                                                                                           | Yes              | Py1                | Yes | Yes                      | Yes | Yes | Yes | no  | Bing Zhang                     |                                |
|                             | 6                                                          | <b>Movie Slice</b><br>(nlsa.movieslice)                           | Extractor                             | Generates movie slice visualization from video files                                                                                  | Yes              | Py1/Py2 not used   | No  | No                       | No  | No  | No  | No  | Sandeep Puthanveetil Satheesan |                                |
|                             | 7                                                          | <b>Person Detector</b><br>(nlsa.video.person_detector)            | Extractor                             | Extract locations of people in an image.                                                                                              | Yes              | Py2                | Yes | Yes<br>(MATLAB external) | Yes | Yes | Yes | Yes | No                             | Sandeep Puthanveetil Satheesan |
|                             | 8                                                          | <b>Person Tracker</b><br>(nlsa.video.person_tracking)             | Extractor                             | Extract locations and paths of people moving in videos.                                                                               | Yes              | Py2                | Yes | Yes<br>(MATLAB external) | Yes | Yes | Yes | Yes | No                             | Sandeep Puthanveetil Satheesan |
| 9                           | <b>Video Analytics Toolbox</b><br>(nlsa.cinemetrics_batch) | Extractor                                                         | Extract shot descriptors from videos. | Yes                                                                                                                                   | Py1/Py2 not used | No                 | No  | No                       | No  | No  | No  | No  | Sandeep Puthanveetil Satheesan |                                |
| Biology, Genomics, Medicine | 1                                                          | <b>FSL</b><br>(mri2mesh)                                          | Extractor                             | Creates a finite element mesh from a set of mr images                                                                                 | Yes              | Py2                | no  | yes                      | yes | no  | no  | no  | Yan Zhao                       |                                |
|                             | 2                                                          | <b>Glomeruli</b><br>(nlsa.msc.diagnosis)                          | Extractor                             | Extract glomeruli from Kidney biopsy images.                                                                                          | Yes              | Py1                | Yes | Yes                      | Yes | Yes | no  | no  | Yan Zhao                       |                                |

## General

| Domain |    | Tool                                       | Type      | Description                                                   | Code | Py1 /Py2 | JSON-LD | DCCKR | Test File | TC  | DPL | DL  | Assignment     |
|--------|----|--------------------------------------------|-----------|---------------------------------------------------------------|------|----------|---------|-------|-----------|-----|-----|-----|----------------|
|        | 1  | <b>Calibre</b> (ebook-converter)           | Converter | Convert e-books to a number of document formats.              | Yes  | N/A      | N/A     | Yes   | Yes       | Yes | No  | Yes | Bing Zhang     |
|        | 2  | <b>CMU Sphinx</b> (ncaa.audio.speech2text) | Extractor | Audio recognition to extract text for speech within audio.    | Yes  | N/A      | Yes     | Yes   | Yes       | Yes | No  | No  | ??who??        |
|        | 3  | <b>Daffodil</b> (daffodil)                 | Converter | Convert formats with a provided DFDL schema to XML.           | Yes  |          |         |       |           | Yes | Yes | Yes | Kenton McHenry |
|        | 4  | <b>DBPedia</b> (ncaa.dbpedia)              | Extractor | Find and define named entities within the given text.         | Yes  | Py1      | Yes     | Yes   | Yes       | Yes | Yes |     | Luigi Marini   |
|        | 5  | <b>FFmpeg</b> (ffmpeg)                     | Converter | Convert between a large number of video formats.              | Yes  | N/A      | N/A     | Yes   | Yes       | Yes | No  | Yes | Bing Zhang     |
|        | 6  | <b>FLAC</b> (flac)                         | Converter | Convert to and from the FLAC format from other audio formats. | Yes  | N/A      | N/A     | Yes   | Yes       | Yes | No  | Yes | Bing Zhang     |
|        | 7  | <b>Ghostscript</b> (ghostscript)           | Converter | Convert between document formats.                             | Yes  | N/A      | N/A     | Yes   | Yes       | Yes | No  | Yes | Bing Zhang     |
|        | 8  | <b>html2doc</b> (html2doc)                 | Converter | Convert HTML to a number of document formats.                 | Yes  | N/A      | N/A     | Yes   | Yes       | Yes | No  | Yes | Bing Zhang     |
|        | 9  | <b>ImageMagick</b> (ImageMagick)           | Converter | Convert between a large number of image formats.              | Yes  | N/A      | N/A     | Yes   | Yes       | Yes | No  | Yes |                |
|        | 10 | <b>ImageMagick</b> (ncaa.image.metadata)   | Extractor | Pull available EXIF image metadata from a given image.        | Yes  | Py2      | Yes     | Yes   | Yes       | Yes | Yes | Yes |                |
|        | 11 | <b>Kabeja</b> (kabeja)                     | Converter | Convert between a handful of 3D and image formats.            | Yes  | N/A      | N/A     | Yes   | Yes       | Yes | Yes | Yes | Kenton McHenry |
|        | 12 | <b>OpenCV - Faces</b> (ncaa.cv.faces)      | Extractor | Find faces in an image and return their locations.            | Yes  | Py1      | Yes     | Yes   | Yes       | Yes | Yes | Yes | Kenton McHenry |

|         |    |                                                           |           |                                                                                                                                     |     |                    |     |     |     |     |     |     |                                   |
|---------|----|-----------------------------------------------------------|-----------|-------------------------------------------------------------------------------------------------------------------------------------|-----|--------------------|-----|-----|-----|-----|-----|-----|-----------------------------------|
| General | 13 | <b>OpenCV - Eyes</b><br>(ncaa.cv.eyes)                    | Extractor | Find eyes in an image and return their locations.                                                                                   | Yes | Py1                | Yes | Yes | Yes | Yes | Yes | Yes | Kenton McHenry                    |
|         | 14 | <b>OpenCV - Closeups</b><br>(ncaa.cv.closeups)            | Extractor | Determine whether an image is a closeup of a person or not.                                                                         | Yes | Py1                | Yes | Yes | Yes | Yes | Yes | Yes | Kenton McHenry                    |
|         | 15 | <b>OpenCV - Profiles</b><br>(ncaa.cv.profiles)            | Extractor | Find human face profiles in an image and return their locations.                                                                    | Yes | Py1                | Yes | Yes | Yes | Yes | Yes | Yes | Kenton McHenry                    |
|         | 16 | <b>Langid</b> (ncaa.nlp.simplelanguage)                   | Extractor | Identify the language of the given text.                                                                                            | Yes | Py2                | Yes | Yes | Yes | Yes | No  | No  | Gregory Jansen                    |
|         | 17 | <b>LibreOffice</b><br>(unoconv)                           | Converter | Convert to and from a variety of document formats.                                                                                  | Yes | N/A                | N/A | Yes | Yes | Yes | No  | Yes | Bing Zhang                        |
|         | 18 | <b>NLTK - Summary</b><br>(ncaa.nlp.simplesummary)         | Extractor | Summarize a body of text.                                                                                                           | Yes | Py2                | Yes | Yes | Yes | Yes | No  | No  | Gregory Jansen                    |
|         | 19 | <b>Siegfried</b><br>(siegfried)                           | Extractor | Extract information about a given file relevant to identifying its type and validating its format.                                  | Yes | Py2                | Yes | Yes | Yes | Yes | No  | No  | Gregory Jansen                    |
|         | 20 | <b>Stanford CoreNLP</b><br>(ncaa.nlp.SNLP)                | Extractor | Natural Language Process extractions such as parts of speech, named entities, language, etc.                                        | Yes | Py2<br>Java equiv. | Yes | Yes | Yes | Yes | No  | No  | Gregory Jansen                    |
|         | 21 | <b>Tesseract</b><br>(ncaa.image.ocr)                      | Extractor | Object Character Recognition (OCR) to extract text from images containing text.                                                     | Yes | Py1                | Yes | Yes | Yes | Yes | Yes | Yes | Bing Zhang                        |
|         | 22 | <b>Tika</b> (ncaa.nlp.tika)                               | Extractor | Document extractions such as language identification, ...                                                                           | Yes | older<br>Py1       | Yes | Yes | Yes | Yes | No  | Yes | Gregory Jansen                    |
|         | 23 | <b>txt2html</b><br>(txt2html)                             | Converter | Convert text documents to HTML.                                                                                                     | Yes | N/A                | N/A | Yes | Yes | Yes | Yes | Yes | Bing Zhang                        |
|         | 24 | <b>Versus - Color Distribution</b><br>(ncaa.versus.image) | Extractor | Generate a distribution of color values within an image to be used for comparing how similar two images are.                        | Yes | Py1                | Yes | Yes | Yes | Yes | Yes | Yes | Bing Zhang                        |
|         | 25 | <b>VLFeat</b> (ncaa.image.caltech101)                     | Extractor | Classify images as to whether they contain objects from the Caltech101 dataset (e.g. people, airplanes, motorcycles, cougars, ...). | Yes | Py2                | yes | yes | yes | Yes | yes | no  | Sandeep Puthanveetil<br>Satheesan |
|         | 26 | <b>Zip</b> (zip)                                          | Converter | Unzip zip archives.                                                                                                                 | Yes | N/A                | N/A | Yes | Yes | Yes | No  | Yes | Bing Zhang                        |

<https://opensource.ncaa.illinois.edu/confluence/display/BD/Transformations> (Under development, Being Refactored)

GeoTiff Extractor by using