GDAL - file management in a virtual directory

Contributed by hk
Monday, 13 October 2014
Last Updated Monday, 13 October 2014

Combine open source software for GIS data management.

GIS data like a set of DEMs or orthophotos is often zipped in one big file (e.g. up to 1 GB) prepared for download.

Not uncommon, only a subset is needed for further GIS analysis. Unzip all files? Not needed, open source software like GDAL and grep offer a solution.

Given a zipped file orthophotos.zip, GIS data is in the subfolder dataset.

```
# list all files in a zipped file by gdal_ls.py

gdal_ls /vsizip/orthophotos.zip/dataset


/vsizip/orthophotos.zip/dataset/data01.TFW

/vsizip/orthophotos.zip/dataset/data01.TIF

/vsizip/orthophotos.zip/dataset/data01.TIF.xml

/vsizip/orthophotos.zip/dataset/data01.TIF.aux.xml

/vsizip/orthophotos.zip/dataset/data02.TFW

/vsizip/orthophotos.zip/dataset/data02.TIF

/vsizip/orthophotos.zip/dataset/data02.TIF.xml

/vsizip/orthophotos.zip/dataset/data02.TIF.aux.xml

/vsizip/orthophotos.zip/dataset/data03.TFW

/vsizip/orthophotos.zip/dataset/data03.TIF

/vsizip/orthophotos.zip/dataset/data03.TIF.xml

/vsizip/orthophotos.zip/dataset/data03.TIF.aux.xml

[...]


# pipe the file list to a text file by gdal_ls, e.g.

gdal_ls /vsizip/orthophotos.zip/data > filelist.txt
```

filter non GIS files by piping grep commands, e.g.

grep -v .xml filelist.txt | grep .TIF > data_to_index.txt

create a raster tile index by gdaltindex e.g.

gdaltindex -src_srs_name sourcesrs mydata.shp --optfile data_to_index.txt

Enjoy your GIS analysis.