Citizen Investigative Journalism
Today Bellingcat releases its latest report, “MH17 – The Open Source Investigation, Two Years Later”, bringing together two years of research by Bellingcat into the fate of Malaysia Airlines Flight 17 on July 17, 2014. As part of this release Bellingcat communicated with experts in various fields for their analysis and opinions of MH17 related...
What went wrong with De WAR?

Having lost 10 years of work, over €100,000 worth of feasibility studies and about €600,000 in crowdfunding one wonders what lessons can be learned...
Where to start?

Existing media are of little help...
When to give up?

Facing the road ahead after a lost battle...

https://www.dezeen.com/2016/09/11/chr
christian-richter-photographs-abandoned-empty-buildings-europe/
Handling paper sources

Tools

hardware  Brother DS-720D: 5ppm duplex mobile scanner  €150
software*  apt-get install scanimage tesseract pdftk gs

Scan a document
$ scanimage --device-name dsseries:usb:0x04F9:0x60E0 \\n  --format=tiff -l 0 -t 0 -x 210 -y 297 >page1.tiff

Convert to searchable pdf using Optical Character Recognition (OCR)
$ tesseract -psm 1 -l nld page1.tiff page1.tiff pdf

Join scanned pages into a single pdf
$ pdftk page1.pdf page2.pdf cat output document.pdf

Reduce the image quality in a pdf
$ gs -sDEVICE=pdfwrite -dCompatibilityLevel=1.4 -dPDFSETTINGS=/ebook \n  -dNOPAUSE -dQUIET -dBATCH -sOutputFile=smalldocument.pdf document.pdf

* open source tools available in most GNU/Linux distributions, tested on Debian
Handling digital sources

Tools
software* apt-get install wget pdfshuffler

Download a website
$ wget -e robots=off --wait 0.25 -r --timestamping --level=inf https://www.domain.tld

Split a combined pdf into separate files
$ pdfshuffler

* open source tools available in most GNU/Linux distributions, tested on Debian
Searching sources

Tools
software* apt-get install recoll poppler-utils antiword pstotext libimage-exiftool-perl \ python-mutagen

Create a search index on a collection of files
create a recoll config file, named /path/to/searchindex/recoll.conf, containing
topdirs = /path/to/files
dbdir = /path/to/searchindex/db
indexstemminglanguages = dutch

create recoll index
$ recollindex -c /path/to/searchindex

Query search index
$ recall -t -c /path/to/searchindex -b -q "query"

* open source tools available in most GNU/Linux distributions, tested on Debian
Connecting the dots

www.debevoegdeinstanties.nl