

- WORKSHOP DIGITAL METHODS -

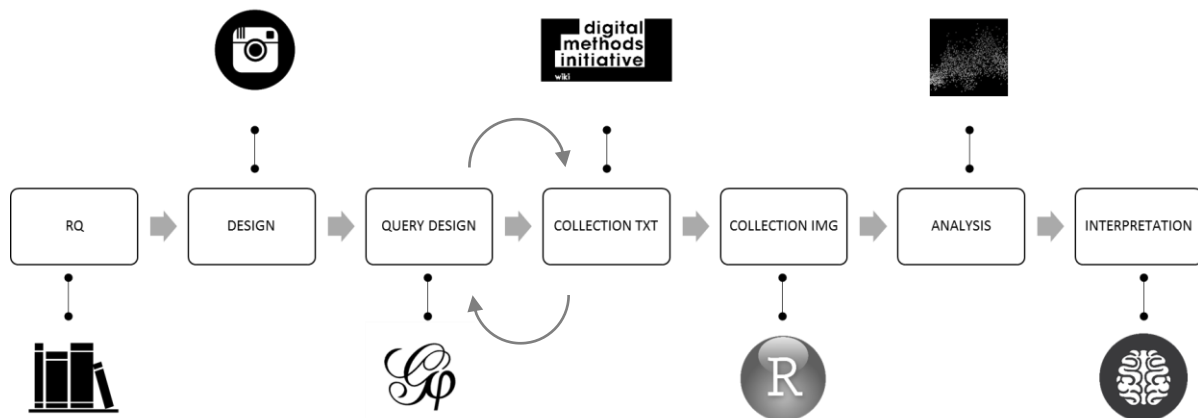
19 NOVEMBER 2015

DIGITAL CULTURE & MEMETIC TREATMENT OF ICONIC NEWS EVENTS

BY: BAS BACCARNE

Dit document vormt een praktisch protocol dat houvast biedt tijdens deze workshop. Ook al doorlopen we de verschillende stappen in groep, hier vind je in grote lijnen terug welke stappen we doorliepen.

1 ONDERZOEKSCOMPONENTEN



2 RQ (& ONDERZOEKSDESIGN)

Definieer RQ

.....

.....

.....

3 QUERY DESIGN

#1

#2

#3

3.1 INITIAL DATA COLLECTION

Ga naar <http://labs.polsys.net/>

Launch de “*Instagram Hashtag Explorer*”, maak verbinding met je Instagram account en geef een hashtag in.

Het startpunt voor deze oefening is een enkele hashtag (onderbouw deze keuze).

Verzamel voor deze eerste fase een **small sample**, zonder preview media.

3.2 CO-HASTAG ANALYSIS IN GEPHI

Download, installeer en open *Gephi* [<http://gephi.github.io/>]

Open de .gdf file (die je in de vorige fase opsloeg) met Gephi

Ga op zoek naar de samenhang door middel van ForceAtlas 2

Links onderaan > Layout > Choose a layout > ForceAtlas 2 > run

(zodra min of meer stabiel: stop)

Maak hastags met een hogere degree beter zichtbaar door de cirkeloppervlakte te vergroten: Rechts bovenaan > Statistics > Run “Average Degree”

Links bovenaan > Diamant logo > Choose a rank parameter > Degree > Apply


Maak clusters in de data duidelijker

Rechts bovenaan > Statistics > Run “Modularity”

Links bovenaan > Partition > Modularity Class > Apply



Interpreteer data door de labels toe te voegen

Rechts onderaan >  > vink autoselect neighbour uit > labels > vink node aan > vink
hide non-selected aan

Filter de data via Rechts Bovenaan > Filters > Attributes > Range > Degree (sleep dit
naar “Queries” er net onder) > speel met de range filter voor een duidelijker beeld
(apply om toe te passen)

Bekijk ook de ruwe data in het *Data Laboratory* (bovenaan)

Doel van deze oefening: query design (zijn er nog hastags die we mee in rekening
moeten nemen voor onze dataverzameling)

4 COLLECTION TXT

Ga eerst naar de Instagram website (instagram.com) en zorg dat je weet hoe groot
de volledige ‘populatie’ is (aantal posts met die hashtag), via de search functie.

posts

Ga terug naar <http://labs.polsys.net/>

Launch opnieuw de “*Instagram Hashtag Explorer*”, maak verbinding met je Instagram
account en geef de eerste hashtag in die je wil scrapen (herhaal deze flow voor elke
mogelijke extra hashtag).

Voor deze oefening willen we wel graag de preview media, user info is optioneel

Open de 2 tabular datafiles & bekijk hoe deze eruit zien (check ook de sample size)

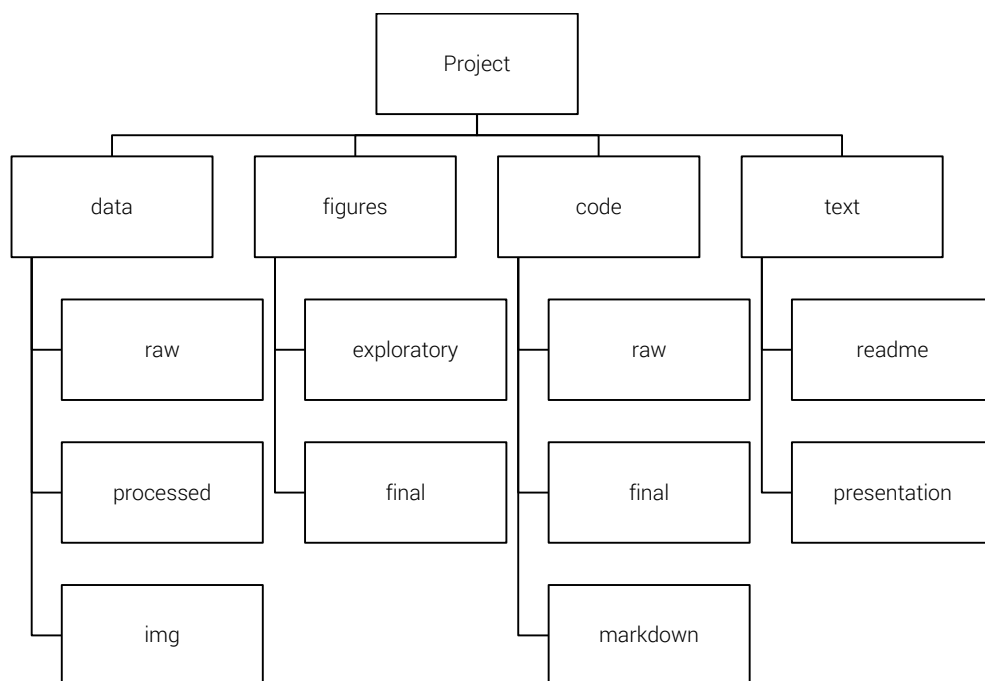
Sla de tabular data op als .txt (tab delimited)

Koppel terug naar je onderzoeksvraag en stuur deze indien nodig bij – bepaal de
variabelen die je nodig hebt voor je analyse (vaak moet je deze nog verwerken om
tot de variabele te komen die je nodig hebt). Pro tip: maak een ‘*tidy data sheet*’. Dit
moet het mogelijk maken de onderzoeksvraag te beantwoorden

TIDY DATA SHEET

Post1					
Post2					
...					

Zorg voor een gestructureerde mappenstructuur (unzip de datafile en plaats de data in data > raw



Een goede mappenstructuur om onderzoeksprojecten in te structureren
(volg deze om geen problemen te hebben met je code)

Geef zoveel mogelijk relevante data mee in je filenames (datum van collectie, query die je gebruikte, sample size) en vermijd spaties, eg:

YYMMDD_query_datatype_samplesize.txt

5 DATA CLEANING & COLLECTION IMG



Download en installeer R Studio [<https://www.rstudio.com/>]

Open R Studio > File > New File > R Script > Sla dit script op in je *code* map (zie schema hierboven)

Ga naar <http://github.com/basbaccarne/MIW15> en kopieer de inhoud van het bestand *masterscript.R* in je eigen R script in R Studio

Zet je cursus onder deze boodschap en doorloop lijn na lijn door ctrl + enter (daarmee voer je deze lijn code uit).

```
18 ▾ #####
19 ▾ # Place cursor below and proceed by pressing ctrl + enter to go through the code
20 ▾ #####
```

Volg de instructies die het script aangeeft (het protocol zit vervlochten tussen code).

Probeer ook te begrijpen wat de code precies doet (daarvoor dient de documentatie die bij de code staat)

Als alles goed ging heb je nu een map met afbeeldingen en een verwerkt

databestand met evenveel rijen als er afbeeldingen in de img map staan. Ga dit zeker na.

6 ANALYSIS IN IMAGEPLOT

Download en unzip ImagePlot [<http://lab.softwarestudies.com/p/imageplot.html>]

Ga naar de map ImageJ en open dit programma

Ga via het menu naar file > open ... > open het bestand *ImagePlot.txt* (in de hoofdmap van ImagePlot) – er opent nu een nieuw venster

Ga in dit nieuw venster naar Macro's > Run Macro

Vink “polar” uit, “save thumbnail images ” aan (enkel de eerste keer), “options” aan en klik ok

Selecteer de txt datafile die uit R Studio kwam en klik ok

Selecteer de folder waar de images staan en klik ok

Kies waar de thumbnails moeten worden opgeslagen (zo kan de volgende analyse die je maakt sneller lopen. Selecteer voor de volgende keer dat je de analyse laat lopen deze nieuwe map als image source folder).

Definieer in welke kolom van de dataset de afbeeldingsnaam staat, welke variabele gebruikt moet worden voor de X-as en welke voor de Y-as en klik ok.

Vink “axes” aan (de rest uit) en klik ok

Stel de instellingen voor je assen in en klik ok.

ImagePlot begint nu te renderen (kan even duren)

Om de afbeelding op te slaan ga je terug naar het hoofdvenster van ImageJ en ga je naar file > save as

Probeer verschillende visualisaties om de onderzoeksvraag te beantwoorden.

Hiervoor moet je terug loskomen van de data en interpreteren (cfr. NVivo) – zie ook 7.1 en 7.2 voor een uitgebreidere analyse.



ANALYSEER, REFLECTEER KRITISCH EN INTERPRETEER

Interessante voorbeelden van dergelijke visualisaties kan je hier vinden:

<https://www.flickr.com/photos/culturevis/>

7 EXTRA

7.1 ANALYSE VAN HET BEELDMATERIAAL

7.1.1 COUNTING THE AMOUNT OF SHAPES IN THE IMAGES

Ga naar het imageJ venster en open ImageShapes.txt (in de map extra's)

Ga naar macros & klik op run macro

Selecteer de folder met images

Kies waar je de output wil opslaan

Het analyseren duurt een tijdje

Deze data kan je of as such analyseren in imageplot (filenames zijn included, verander gewoon de extensie naar .txt) of toevoegen aan je oorspronkelijke datafile (combinatie met e.g. timestamp data, likes, comments, ... mogelijk)

7.1.2 BRIGHTSNESS/HUE/SATURATION OF THE IMAGES

Ga naar het imageJ venster en open ImageMeasure.txt (in de map extra's)

Ga naar macros & klik op run macro

Selecteer de folder met images

Kies waar je de output wil opslaan

Het analyseren duurt een tijdje

Deze data kan je of as such analyseren in imageplot (filenames zijn included) of toevoegen aan je oorspronkelijke datafile (combinatie met e.g. timestamp data mogelijk)