*Algemeen: dit zijn niet echt "uitwerkingen", maar vooral wat algemene aanwijzingen en commentaar bij het uitvoeren van de oefeningen.*

**Opdracht bronnen voor boeken, artikelen**

1.a Je ziet dat Google Books echt in de inhoud van de boeken zoekt. Je krijgt dus ook boeken te zien waar de gevraagde termen (hier de naam van Fortuyn) alleen een keer ergens op een bepaalde bladzijde voorkomen. Die bladzijde(n) krijg je (als auteursrecht het toestaat) ook te zien. Binnen een gevonden boek kun je ook nog full-text verder zoeken op eventueel nieuwe zoektermen.

In Hathitrust zitten boeken die bijna allemaal ook in Google Books zitten. Ook daarin kun ej full-text zoeken, maarmet een wat ander zoekinterface. Ook daar kun je binnen een boek verder zoeken.

Worldcat geeft je alleen catalogus-records van boeken. En zoekt dus ook alleen in de daarin aanwezige metadata. Je vindt hier dus ook veel minder boeken, omdat ze echt (helemaal) **over** Pim Fortuyn moeten gaan (of door hem geschreven moetne zijn), voordat die naam in de metadata voorkomt.

Amazon zoekt in de metadata van de boeken die ze verkopen. Behalve informatie over het boek (ook over bestellen) vind je vaak ook uitgebreider beschrijvingen en "recensies" door andere gebruikers. Ook hier moeten boeken echt (vrijwel) helemaal over Fortuyn gaan om zo gevonden te worden.

In Libraything vindt je eveneens metadata van boeken, aangevuld met gegevens die gebruikers hebben toegevoegd, zoals besprekingen en beschrijvingen. Ook hier zoek je alleen in de metadata (en eventueel die extra beschrijvingen).

In Delpher vindt je op *“willem van oranje”*  936 boeken. Net als bij Google wordt hier ook full-text in de geOCR-de tekst gezocht en kun je de pagina's met de zoektermen meteen te zien krijgen (en kun je ook door de hele boeken bladeren - als je geduld hebt ....)

1.b Op de zoekvraag *relationship musicality intelligence* vindt je in Scholar ruim 7000 resultaten. Allemaal wetenschappelijke artikelen of (een enkele keer) boeken. Je kunt dit inperken met Google-syntax, door te zoeken op een exacte phrase met wildcards "*relationship \* musicality \* intelligence"*  (of de echte exacte phrase "*relationship between musicality and intelligence"*) die 34 (resp. 39 !?) resultaten oplevert. Je kunt ook proberen te eisen dat bepaalde zoekwoorden in de titel moeten voorkomen. Zo levert *relationship musicality intitle:intelligence*  62 resultaten, idem met alleen "musicality" in de titel 68 resultaten en met alleen "relationship" in de titel 46 resultaten. Gezien deze aantallen en de vermoedelijke relevantie van het gevondene is het niet zinvol te eisen dat minstens 2 van die zoektermen in de titels moeten voorkomen.

In sEURch levert de oorspronkelijke zoekvraag 1820 resultaten op. Ook tijdschriftartikelen en (relatief wat meer) boeken. Met de exacte phrase, waarop je hier op dezelfde manier blijkt te kunnen zoeken, houdt je er daarvan nog 9 over. Er is niet te vinden hoe eventueel alleen op titelwoorden te zoeken. Bij de filters links lijken er geen te zijn die het resultaat nog zinvol zouden kunnen inperken.

In BASE levert de oorspronkelijke zoekvraag 637 resultaten, bijna allemaal artikelen. Alleen op titels zoeken heeft geen zin, want dan moeten alle zoektermen in de titel staan en blijft erniets over. De tekst "entire document" als aanduiding waarin je zoekt, betreft het gehele "metadata-document" (d.w.z. alle velden daarvan), en dus niet de volledige tekst van de primaire documenten zelf!

**Nieuws**

2. Zoek hier in eerste instantie alleen op de beslist noodzakelijke kernwoorden *obama “intellectual property” tpp* . Dat levert eigenlijk meteen al relevante resultaten. Woorden als “paragraph”, “president”, “barack” of “treaty” mag je weglaten.

In de NL Google Nieuws krijg je met de Engelse term “intellectual property” toch weer berichten uit Amerikaanse bladen. Je kunt dat via de “search tools’ wel inperken tot alleen maar Nederlandstalig, maar dan blijft er (niet echt onverwacht) bijna niets over. Met “auteursrecht” komt er wel wat uit, maar het merendeel daarvan gaat eigenlijk over de komende verliezingen, waar bij het TPP-verdrag maar zijdelings even genoemd wordt.

Op de site van de BBC blijk je met de summiere zoekvraag zoals bij Google, in eerste instantie vooral heel oude berichten te vinden. Je kunt dan op "News" inperken, maar ook dat mag niet baten. Er is ook geen optie om op datum te sorteren.

Op nu.nl blijkt geen van de zoekvragen die ik kon bedenken (ook alleen Obama + TPP of Obama + handelsverdrag) iets op te leveren.

**Oude informatie**

3.a De Way-back-machine spreekt redelijk voor zich zelf. Alleen moet je weten dat als je een jaar (waarin copieën gemaakt zijn) hebt aangeklikt, je ook nog in de dan verschijnende kalender een dag met een blauw bolletje (een dag waarop een copie gemaakt is) moet aanklikken.

3.b. In Delpher zou ik het jaartal niet in de zoekactie zelf betrekken, omdat in die tijd natuurlijk niet expliciet gesteld hoefde te worden dat het om de spelen van 1928 ging. Via "Uitgebreid zoeken" kun je echter wel mooi op verschijningsdatum van de kranten inperken, dus bijvoorbeel zoeken in kranten die tussen pakweg januari 1927 en december 1929 zijn uitgekomen. Met *financiering olympische spelen amsterdam* vind je dan 13 artikelen, die echter niet allemaal relevant lijken (al is dat uit de resultatenlijst niet altijd meteen op te maken - pas als je enige tijd met je muis boven een resultaat blift hangen, verschijnen er enkele regels tekst uit het gevonden artikel). Anderzijds mag je misschien "Amsterdam" ook wel weglaten, want dat was ook al tamelijk impliciet als men in die periode over de spelen schreef. Daarmee maak je er 17 van. Met "fondsen" (en dan moet je die tijdinperking weer opnieuw invoeren :( )vind je 69 resultaten. Maar ook hier lang niet altijd relevant (soms komt het woord "fondsen"zelfs niet in hetzelfde artikel voor als de "olympische spelen".

**Blogs, Feeds**

4.a *privacy adobe e-book reader* levert in Icerocket 6 blogposts. Met *ebook* (zonder streepje) zijn dat er 9. Rechts van de titels zie je telkens hoeveel dagen ze oud zijn. Kennelijk krijg je automatisch al de recentste bovenaan te zien. In het geavanceerde zoekscherm kun je nog wel wat extra zoekopties gebruiken, bijvoorbeeld een zoekterm tot de titel van een blogpost beperken; en ook de link optie. Gevonden resultaten blijken echter nauwelijks relevant te zijn. Kennelijk speelt dit onderwerp al niet meer.

Twingly geeft op die zelfde vraag 23, resp. 55 resultaten. Daar moet je expliciet aangeven dat je ze niet gerankt, maar op datum gesorteerd wilt hebben. Een geavanceerd zoekscherm is daar verder niet.

Google (via de sluiproute die in de opdracht gegeven werd) levert ruim 27000 resultaten. De relevantiesortering moet je in sortering op datum veranderen. Met de meer gespecificeerde vraag *privacy concerns Adobe Digital Editions e-book reader* bliijven er daar nog 338 van over. Maar daarbij blijkt ook wel erg ruim op synoniemen gezocht te zijn. Als je alle termen letterlijk laat zoeken, blijft er nog maar 1 over. Maar die blijkt niet relevant.

4.b In CTRQ vind je met *nrc wetenschap* feeds (alle afzonderlijke afleveringen) waarin de woorden NRC en wetenschap voorkomen; met *nrc wetenschap site:nrc.nl* vind je de feeds (maar nog altijd alle afzonderllijke afleveringen) van de NRC zelf.

In RSSsearchhub vind je alleen feeds als geheel. De zoekvraag *nrc wetenschap* is genoeg om 4 wetenschapsfeeds te vinden, waarvan 3 van de NRC zelf.

**Twitter , Social web, ...**

5.a Aan het voorbeeld *uitleningen ebooks* zie je makkelijk dat het bij Twitter verschil maakt of je op enkel- of meervoud zoekt en of je al dan niet een koppelstreepje in e-books gebruikt. Filter op positieve of negatieve berichten kijkt alleen of er ☺ of ☹ in het bericht gebruikt is.

Topsy sorteert anders dan Twitter, in eerste instantie op relevantie in plaats van op recentheid. Je kunt dat wel anders kiezen. Jammer genoeg zie je niet meteen hoeveel er gevonden is. Je ziet dat enkel- of meervoud en al dan niet koppelstreepje ook daar verschil maakt. Geavanceerd zoeken (onder het tandwieltje) biedt minder opties dan twitter zelf.

In Social-searcher krijg je de zoekresultaten in Facebook, Twitter en Google+ naast elkaar te zien. Helaas zijn de resultaten uit die eerste twee alleen tot een heel recente periode beperkt. Bij onderwerpen die op dit moment niet erg leven (zoals die uitleningen), vind je daar dus soms niets.

In SocialMention levert de vraag over de ebooks wel meteen redelijk wat op, maar dat komt omdat ook in (o.a.) blogs wordt gezocht. Bij inperking op "microblogs"(i.e. Twitter) blijft er ook niets over.

5.b Bij Facebook is recent iets veranderd, waardoor de gestructureerde zoekacties niet altijd meer zo makkelijk werken. Als je bijvoorbeeld intikt *favorite movies of people living in utrecht* dan maakt Facebook daar automatisch van *favorite movies of current utrecht residents*. Dat levert aanvankelijk 2 en naa doorklikken een veel langere lijst films op.

Van *books liked by people working for microsoft* maakt Facebook *Favorite books of current Microsoft employees.* Ook hier aanvankelijk weer 2 en vervolgens een hele lijst.

**Zoeken naar images en foto's**

6.a Zoeken op alleen de achternaam van Einstein (zonder Albert) in combinatie met het woord "tong", en dan liefst in het Engels (tongue) omdat de foto op veel meer Engelstalige sites staat, geeft een overvloed aan exemplaren van de bedoelde foto en allerlei varianten en bewerkingen ervan. Tegenwoordig is het in veel systemen (en zeker in Google) moeilijk - zo niet onmogelijk - te zien waar de woorden Einstein en Tongue voorkwamen waarop de foto's zijn gevonden. Vroeger was vaak wel eenvoudig zichtbaar dat een of twee van die zoektermen in de naam van de imagefiles voorkwamen.

6,b Met simpele zoekvragen als *arrow up* (ook weer in het Engels omdat je dan een veel grotere variëteit aan resultaten krijgt dan met NL termen) krijg je een heleboel voorbeelden. In Google vind je onder de Search Tools mogelijkheden om in te perken op grijstinten of om ruwweg op enkele resoluties in te perken. Op Google vind je vaak veel meer dan op gespecialiseerde clip-art sites, zij het dat dergelijke afbeeldingen daar vaak in systematieken zijn ondergebracht, zodat je niet over zoekwoorden hoeft na te denken.

6.c Ook hier gewoon de Search Tools uitproberen

6.d Je zult merken dat het voor duidelijk herkenbare objecten op foto's vaak beter gaat dan voor bijvoorbeeld zwart-wit tekeningen van een cartoon als Fokke&Sukke, omdat voor een computer al die zwart-wit tekeningen al heel erg op elkaar lijken.

6.e Gewoon uitproberen

6.f Gewoon uitproberen. Voor creative commons materiaal bieden Flickr en Google filters.

6.g Met *obama marijuana dangerous alcohol* vind je in Goo(in he-video, YouTube en Blinkx wel meteen uitzendingen waar het over deze uitspraak gaat, maar het is dan nog wel even zoeken om de woorden van Obama zelf te vinden. In elk geval kom je niet automatisch meteen bij het fragment waar de gezochte tekst wordt uitgeproken. In Voxalead kies je via het vlaggetje in de zoekbalk de gewenste taal van de video's. In het Engelstalige nieuws vind je twee fragmenten waarin het (op NBC News) over die uitspraak van Obama gaat, van resp, 20 en 19 januari 2014. Ook daar niet direct Obama zelf aan het woord, maar je kunt wel meteen in de getoonde stukjes transcriptie het fragment kiezen waar die woorden (door een TV-presentator) gesproken worden.

Met *Dijsselbloem Cyprus* vind je in Voxalead ongetwijfeld (ook) een video waarin die Cyprus-commotie ter sprake komt, maar nog niet meteen het fragment waar het om die "commotie" gaat. Met dat woord als zoekterm erbij kom je inderdaad ook meteen bij het gewenste fragment, al na 14 seconden in het Journaal van 7 uur op 26 maart 2013.

6.h In het BBC archief is die specifieke uitzending met de eigen zoekfunctie niet goed te vinden. Gewoon in Google zoeken naar *Panorama Thatcher 1970 site:bbc.co.uk* levert wel meteen als eerste de betreffende pagina in het BBC-archief (zelfs 3x).

6.i Gewoon uitproberen.

**Zoeken naar naslag en antwoorden**

7. In de NL-wikipedia is het niet handig om te zoeken op de combinatie *tinnitus voorkomen* omdat bij het zoeken geen onderscheid gemaakt kan worden tussen *vóórkomen* en *voorkómen .* Alternatief is om op *tinnitus preventie* te zoeken. Dat levert het algemene lemma over tinnitus (1474 woorden) met een paragraafje van 4 regels over preventie, maar niet specifiek voor musici. In het ook gevonden lemma over hyperacusis gat het meer om gewenning aan oorsuizen dan over preventie, en al helemaal niet over musici. Op *tinnitus musici* vind je wel drie resultaten, waaronder het lemma over "gehoorschade" (585 woorden) de meest relevante informatie bevat. De combinatie *oorsuizen musici* levert overigens vrijwel hetzelfde resultaat.

In de EN-wikipedia vind je op de combinatie *tinnitus prevention musicians* al 3 resultaten, maar geen daarvan geeft direct antwoord op de vraag. Het lemma *Tinnitus* (5351 woorden - veel meer dan in de NL-versie) is daar ook niet meteen bij. Daarin wordt ook niet specifiek ingegaan op musici. In de korte paragraaf over "prevention" wordt wel verwezen naar *Ear Plugs*. In dat lemma (*Earplug* - 4389 woorden) blijkt het wel over Tinnitus en musici te gaan. Met de oorspronkelijke zoekvraag vond je dit echter niet omdat alleen de werkwoordsvormen "prevent" en "preventing" voorkomen.

Door in plaats van Wikipedia's eigen zoekfunctie, Google te gebruiken (die wel op woordstammen zoekt) had je dit wel meteen gevonden: *tinnitus prevention musicians site:wikipedia.org*

In Google als geheel vind je met *tinnitus prevention musicians* een grote hoeveelheid informatie waaronder informatie van een waarschijnlijk betrouwbare organisatie als de "American Tinnitus Association" en een PDF-document van 100 blz dat helemaal over het voorkomen van oorschade bij musici gaat.

In Yahoo-answers vind je met *tinnitus prevention musicians* geen resultaten, maar met *tinnitus protection musicians* wel een groot aantal vragen, met daarbij de gegeven (praktijk-)antwoorden, soms 1 antwoord, soms meer per vraag.

In Quora vind je met *tinnitus musicians* een aantal gestelde vragen. De eerste daarvan (How can musicians prevent tinnitus?) was overigens enige jaren geleden als test door mijzelf gesteld, maar heeft intussen wel 4 antwoorden gekregen. Maar ook bij de verdere vragen die dit opleverde zitten er nog wel een paar die relevant zijn. Dat is zeker het geval als je de zoekvraag meteen al volledig specificeert *tinnitus musicians prevention.* Je ziet aan het resultaat daarvan trouwens dat Quora wel word-stemming toepast (prevent, preventative) , dit in tegenstelling tot Wikipedia.

**Zoeken naar feiten en data**

8.a With *population rio de janeiro , unemployment netherlands , unemployment greece* krijg je snel die betreffende gegevens te zien

De vragen waren geïnterpreteeerd als resp: "*Rio de Janeiro, Brazil" " city population"* en, *netherlands/greece "unemployment rate" ,*  Door in het pop-up icoontje van een open boek te klikken, zie je dat de gegevens afkomstig waren uit een eigen Knowledge Base en een veelheid aan externe bronnen.

Met bijv. *education suriname vs indonesia* of *comparison education suriname indonesia* vind je vergelijkende cijfers. Hoewel die vraag geïnterpreteerd wordt als *Suriname/Indonesia students*  krijg je cijfers over aantallen deelnemers aan diverse soorten onderwijs.

8.b. In Zanran vind je met *milk exports* meteen al een heleboel resultaten. Door met de muis over de kleine blokjes voor de resutlaten te gaan, krijg je een preview van de tabel, grafiek of hele pagina op basis waarvan Zanran dit gegeven geselecteerd heeft. Dat maakt het makkelijker om te selecteren welke gegevens je echt wilt bekijken.

Filtermogelijkheden (in de grijze balk) zijn er op landen/domeinen, op datum (recentheid) en op filetype. Filteren op Excel-sheets kan nuttig zijn als je de gegevens zelf ook meteen weer in een spreadsheet zou willen verwerken (maar als gewenste gegevens alleen maar in andere file-formaten blijken voor te komen, zul je het daar natuurlijk mee moeten doen).

Als je specifiek naar Excel-sheets zoekt kan dat natuurlijk ook met Google door:

*milk exports filetype:xls*

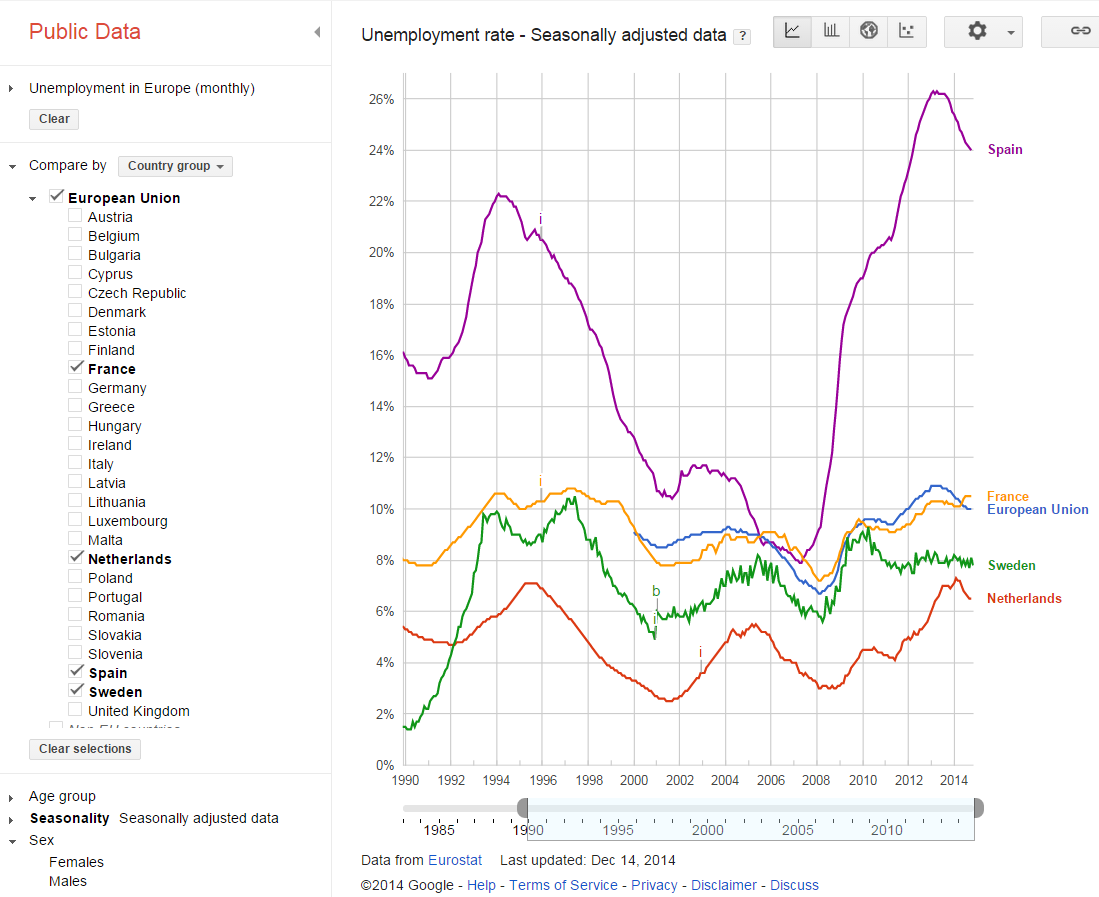
Dat levert nog veel meer resultaten, maar geen preview mogelijkheid.

Met *pig meat export netherlands* krijg je 2718 resultaten. *export* of *exports* blijkt geen verschil te maken. Ook maakt het geen verschil of je een AND-operator tussen zoekwoorden zet. Met "pig meat export" zoek je op een phrase, waarbij - anders dan bij Google - wel nog op de woordstammen van woorden uit de phrase gezocht wordt (dus ook op o.a. *exports , exporters* ).

Met *rice production vietnam* vind je veel dat ook echt over de productiecijfers van rijst gaat. Toch zou je het kunnen verfijnen met de exacte phrase *"rice production"* (al moet je je realiseren dat je dan resultaten mist waarin bijvoorbeeld sprake is van *"production of rice"* ).

Met *Infant mortality rate India 2010* vindt Zanran meer dan 11.000 resultaten, maar bij de eerste ziten meteen al een aantal verschillende overzichten waar die cijfers in staan, soms nog uitgesplitst naar de Indiase deelstaten.

8.c De unemployment rate uit Eurostat levert meteen dit soort mooie, interactief bij te stellen grafiekjes op. Je kunt ook (beperkt) op leeftijdsgroep en op sexe uitsplitsen.



8.e Met *emission greenhouse* gases vind je hier 2174 “documenten”. Daarvoor was geen AND-operator nodig. Bij het zoekresultaat zitten ook gewoon artikelen. Onder “Resource Types” kun je filteren op datasets, Dat zijn er 1778.

Door te klikken op een titel kom je eerst nog terecht op een tussenpagina van Datacite zelf, met nadere gegevens (metadata) over de dataset. Via de DOI-link (digital object identifiers worden niet alleen gebruikt voor artikelen, maar ook om permanent naar dit soort gegevens te linken) kom je op een “landing page” waar je de dataset(s) kunt bekijken of downloaden. Overigens kom je daarbij ook nog resultaten tegen die alleen de tekst van een artikel of proefschrift blijken te zijn, zonder direct herkenbare “data”.