

HHS - HBO-ICT - RTT

-

Topic

Data Driven Society

-

1. Casus De Tweede Kamer
2. Casus Gemeente Den Haag & Innolab
3. Casus CBS
4. Casus Lokale overheid
5. Casus Beeld & Geluid

Fenneke Mink MA

HHS Docent HBO-ICT

April 2018

# Inhoudsopgave

<b>1 Data Driven Society: Topic – Outline .....</b>	<b>3</b>
<b>2 Casus De Tweede Kamer.....</b>	<b>4</b>
<i>Casus 1 Open Data voor de Tweede Kamer: .....</i>	<i>6</i>
<i>Casus 2 Applicaties voor de Tweede Kamer .....</i>	<i>7</i>
<i>Casus 3 Controleer de Tweede Kamer.....</i>	<i>8</i>
<b>3 Casus Innolab.....</b>	<b>10</b>
<b>4 Casus CBS Data voor Wikipedia .....</b>	<b>13</b>
<b>5 Casus Open Data voor de lokale overheid .....</b>	<b>15</b>
<i>Opdrachtgever Gemeente Den Haag en Provincie Zuid – Holland .....</i>	<i>15</i>
Onderzoeksthema's voor de Gemeente Den - Haag: .....	15
Onderzoeksthema's voor Provincie Zuid - Holland:.....	15
Opdracht.....	16
<b>6 Casus Beeld en Geluid .....</b>	<b>17</b>
Inleiding .....	17
Opdracht.....	17
Over Beeld en Geluid.....	18
<b>7 Data Driven Society: mogelijke bronnen .....</b>	<b>19</b>
<b>8 Data Driven Society: mogelijke deelvragen .....</b>	<b>21</b>
<i>Overheid open Data.....</i>	<i>21</i>
<i>(Community) platform.....</i>	<i>21</i>
<i>Semantic web .....</i>	<i>21</i>
<i>Geo-informatie .....</i>	<i>22</i>
<i>Verandermanagement ten opzichte van Open Data in overheid.....</i>	<i>22</i>
<i>Linked Open Data .....</i>	<i>22</i>

# 1 Data Driven Society: Topic – Outline

Overheden zijn enkele jaren actief in het aanbieden van grote databestanden in het kader van transparantie. Hiernaast om andere overheden en ontwikkelaars in de gelegenheid te stellen om interessante applicaties te kunnen ontwikkelen voor het grote publiek. Nationale overheden, zoals die van Amerika, Engeland en Nederland, maar ook lokale overheden zoals provincies en gemeenten hebben de afgelopen jaren al vele projecten opgezet om grote databestanden online te ontsluiten.

Hoewel het eenvoudig lijkt om databestanden 'op te pakken en online te zetten', ligt daar een heel proces aan ten grondslag. Naast het checken van de technische randvoorwaarden spelen verschillende zaken mee voordat data online gezet kan worden, denk bijvoorbeeld aan de validiteit van de data, de actualiteit van de data, de bruikbaarheid van de data en de verantwoordelijke persoon/dienst/organisatie voor de data in termen van privacy en security.

De overheid en culturele instellingen hebben een grote hoeveelheid databestanden online beschikbaar. En zij willen in de nabije toekomst graag nog meer databestanden openstellen en koppelen ten behoeve van de maatschappij. Hierbij willen zij graag door de studenten HBO-ICT, geholpen worden. De studenten treden op als onderzoeker in opdracht van verschillende opdrachtgevers. Zij doen dit in een groepje waarin zij een vraagstuk betreft (Linked) Open Data onderzoeken. Op basis van de onderzoeksresultaten ontwikkelen zij een advies aan de opdrachtgever. De studenten weten antwoorden te geven op de vragen van de opdrachtgever en haar doelgroep in een bepaalde situatie. Ten grondslag aan het onderzoek ligt een degelijk onderzoeksrapport en posterpresentatie waarin de studenten de opgedane inzichten presenteren en adviseren over de te nemen vervolg stappen.

In de volgende hoofdstukken worden de opdrachten van de verschillende opdrachtgevers weergegeven.

Let op! Het opgeven van de keuze voor een opdracht per groepje vindt plaats in week 1. De groepjes die voor dit topic kiezen ontvangen hierover bericht per email.

## 2 Casus De Tweede Kamer

### Aanleiding

De Koninklijke Bibliotheek werkte onlangs aan een project om in krantenartikelen namen te herkennen (named entity recognition), om die vervolgens te identificeren en te linken aan (bijvoorbeeld) DBpedia-beschrijvingen. Op die manier kan niet alleen worden gezocht op entiteiten uit datasets als DBpedia, Wikidata, VIAF, maar kunnen er ook zoekmogelijkheden ontwikkeld worden op eigenschappen van die entiteiten. Een voorbeeld van deze technologie is te vinden in het KB Lab in de xportal<sup>1</sup>.

### Tweede Kamer

De komende tijd onderzoekt de Tweede Kamer de mogelijkheden van een Open Data Portaal aan de hand van named entity recognition. Zij onderzoeken de mogelijkheden van het koppelen van data als hierboven beschreven en hebben hier advies bij nodig, een lab omgeving met een werkende testomgeving zou hierbij helpen. In de testomgeving kan bijvoorbeeld gezocht worden op door communities gecontroleerde entiteiten in de parlementaire data en waar er gezocht kan worden op bepaalde eigenschappen van deze entiteiten (semantisch zoeken). Er wordt dan niet alleen zoeken op Frans Timmermans<sup>2</sup> maar ook op alle entiteiten (personen) geboren in Maastricht<sup>3</sup>.

### Van open naar linked data

Nadat de afgelopen jaren steeds meer overheidsinformatie is “geopend” voor haar burgers wordt het hoogtij om de informatie te koppelen om tot zinvolle nieuwe inzichten te komen. De komende weken onderzoeken de studenten hoe Open Data en Linked Open Datatechnieken een bijdrage kunnen leveren voor het “openen” van informatie van de Tweede Kamer. De studenten zoeken zelf naar relevante en geschikte datasets om te combineren met data van het Open Data Portaal van de Tweede Kamer<sup>4</sup>. Zij onderzoeken welke combinaties van data en technieken resulteren in de meest zinvolle inzichten.

---

<sup>1</sup> <http://kbresearch.nl/xportal>.

<sup>2</sup> (<https://www.wikidata.org/wiki/Q32681> of [https://www.parlement.com/id/vg09llk9rzrp/f\\_c\\_g\\_m\\_frans\\_timmermans](https://www.parlement.com/id/vg09llk9rzrp/f_c_g_m_frans_timmermans))

<sup>3</sup> (<https://www.wikidata.org/wiki/Q1309>).

<sup>4</sup> <https://opendata.tweedekamer.nl/>

De centrale vraag van de Tweede Kamer is:

- *“Welke rol kan Linked Open Data spelen in het toegankelijk maken van informatie van de Tweede Kamer?”*

En mogelijke deelvragen waar zij antwoord op hopen te vinden zijn:

- Welke datasets zijn, gezien onze data, het meest geschikt (dekkingsgraad)?
- Welke datasets zijn het beste bruikbaar (inhoudelijk)?
- Vragen rond eenduidige, stabiele identificatie van entiteiten.
- Wat zijn de mogelijkheden als gebruik wordt gemaakt van de API van het opendataportaal?
- Is het mogelijk een proeftuinomgeving te bouwen waar getest kan worden met het automatisch herkennen van entiteiten?
- Wat zijn mogelijkheden van crowdsourcing om de data verder te vervolmaken?

### **Beoogd eindresultaat**

De Tweede Kamer beoogd advies te ontvangen over ten minste;

- Het koppelen van entiteiten, i.h.b. personen (indieners, sprekers) in de data van de Tweede Kamer aan externe open datasets als DBpedia, Wikidata, Parlement.com en/of VIAF.
- Het koppelen van externe datasets. Op die manier kunnen (externe) open data gekoppeld worden aan data/entiteiten uit het GGM/opendataportaal.

### **Begeleiding**

Medewerkers van de Tweede Kamer zijn alleen beschikbaar voor studenten via docent:

Fenneke Mink

## Casus 1 Open Data voor de Tweede Kamer:

Onderzoek naar gebruik opendata van de Tweede Kamer

### **Aanleiding<sup>5</sup>**

De Tweede Kamer is een van de meest transparante overheden van Europa. Parlementaire informatie wordt via verschillende informatiekkanalen beschikbaar gesteld. Denk hierbij aan [www.tweedekamer.nl](http://www.tweedekamer.nl), [www.statengeneraaldigitaal.nl](http://www.statengeneraaldigitaal.nl) en <http://data.overheid.nl><sup>6</sup> (open dataportaal). Het is voor de Tweede Kamer momenteel niet altijd zichtbaar welke (overheids)organisaties en applicaties gebruik maken van de TK-opendata, die zij beschikbaar stelt via het dataportaal.

### **Opdrachtschrijving**

Onderzoek welke (overheids)organisaties gebruik maken (en voor welke doeleinden) van TK-opendata en hoeverre dit wordt gebruikt voor het ontwikkelen van applicaties. Inventariseer hiernaast in hoeverre TK-opendata bruikbaar is voor deze organisaties en wat verbeterpunten zijn. Het resultaat van dit onderzoek is inzicht in het gebruik, eisen en wensen aangaande TK-opendata en advies om het TK-opendata aanbod te kunnen optimaliseren.

### **Concrete invulling**

Uitvoeren van een casestudie (kwantitatief onderzoek) naar gebruik en bruikbaarheid van TK-opendata.

### **Aandachtspunten / onderzoeksvragen**

- Welke (overheids)organisaties maken gebruik van TK-opendata?
- Voor welke doeleinden wordt TK-opendata gebruikt?
- Welke applicaties maken gebruik van TK-opendata?
- In hoeverre is TK-opendata bruikbaar? Welke eisen/wensen zijn er rondom de beschikbaar gestelde dataset?
- Wat zijn mogelijk andere applicaties voor gebruik van TK-opendata?

---

<sup>5</sup> Bron: Opdrachtvoorstel geformuleerd door dhr. E Driessen.

<sup>6</sup> Dit dataportaal van de Rijksoverheid maakt datasets van alle Nederlandse overheden vindbaar.

**Wat levert dit op:**

- Inzicht in gebruik en bruikbaarheid van TK-opendata.
- Advies om aanbod TK-opendata te kunnen optimaliseren.
- Advies in mogelijk nieuwe toepassingen die gebruik kunnen maken van TK-opendata.

**Casus 2 Applicaties voor de Tweede Kamer**

Onderzoek naar voor welke doeleinden het interessant kan zijn om TK-opendata te gebruiken en welke applicaties hierop kan worden (door)ontwikkeld.

**Aanleiding<sup>7</sup>**

Alle Nederlandse overheidsinstellingen bieden hun dataset aan op <http://data.overheid.nl>. Diverse (overheids)instellingen maken gebruik van deze datasets om hun dienstverlening aan burgers en ondernemers te verbeteren.

**Opdrachtomschrijving**

Onderzoek voor welke doeleinden het interessant kan zijn om TK-opendata te gebruiken en welke nieuwe applicaties met de dataset van kan worden (door)ontwikkeld. En breng hierover advies uit aan de Tweede Kamer.nl. Hierbij mag je gebruik maken om diverse open-datasets met elkaar te combineren.

**Concrete invulling**

Uitvoeren onderzoek naar bruikbaarheid van TK-opendata voor betere dienstverlening aan burgers en ondernemers.

**Aandachtspunten / onderzoeksvragen**

- Voor welke doeleinden wordt TK-opendata momenteel gebruikt?
- Welke applicaties maken nu al gebruik van TK-opendata?
- Welke (nieuwe) applicaties kunnen ontwikkeld worden met de dataset die Tweede Kamer beschikbaar stelt? En welke meerwaarde kunnen deze applicaties hebben voor burgers en ondernemers?

---

<sup>7</sup> Bron: Opdrachtvoorstel geformuleerd door dhr. E Driessen.

**Wat levert dit op:**

- Inzicht in voor welke toepassingen TK-opendata gebruikt wordt en welke nieuwe applicaties met TK-opendata mogelijk is.
- Advies in mogelijk nieuwe toepassingen die gebruik kunnen maken van TK-opendata.

**Casus 3 Controleer de Tweede Kamer**

Onderzoek in hoeverre de openbare datasets op opendata portaal door Tweede Kamerleden gebruikt kan worden om de regering te kunnen controleren.

**Aanleiding<sup>8</sup>**

Tweede Kamerleden hebben als volksvertegenwoordigers een controleren taak van de regering. Jaarlijks verantwoord de regering de overheidsuitgaven aan de Tweede Kamer.

Kamerleden maken momenteel geen optimaal gebruik van de beschikbare data die via het opendata portaal van de overheid (<http://data.overheid.nl>) wordt beschikbaar gesteld. Het is interessant om te onderzoeken in hoeverre deze datasets gebruikt kan worden door Kamerleden om haar controlerende taak (aangaande de uitgaven van de overheid) van de regering uit te kunnen voeren.

**Opdrachtschrijving**

Onderzoek welke datasets die via het open dataportaal beschikbaar worden gesteld bruikbaar kan zijn voor Kamerleden om haar controlerende taak van de regering uit te kunnen voeren. Het gaat hierbij om de controle van de uitgaven van de overheid. Adviseer hierbij hoe de datasets gebruik kan worden.

**Concrete invulling**

Uitvoeren onderzoek naar de datasets die gebruikt kunnen worden voor Kamerleden om hun controlerende taak van de regering uit te kunnen voeren. Realiseer een datamodel (in Excel) en breng advies uit hoe dit als dienst vanuit Dienst Informatie en Archief kan worden aangeboden.

---

<sup>8</sup> Bron: Opdrachtvoorstel geformuleerd door dhr. E Driessen.



**Aandachtspunten / onderzoeksvragen**

- Welke overheidsorganisaties bieden financiële (verantwoordings)informatie op het open dataportaal?
- Welke datasets kunnen interessant zijn om Kamerleden van informatie te voorzien voor hun controleren taak van regering?
- Welk datamodel kan in Excel hiervoor gemaakt worden?
- Hoe kan Dienst Informatie en Archief deze datasets aanbieden als dienst aan Kamerleden?

**Wat levert dit op:**

- Inzicht in bruikbaarheid van data die het opendata portaal biedt, in het bijzonder voor gebruik door Kamerleden.
- Advies aan Dienst Informatie en Archief om haar dienstverlening aan Kamerleden mogelijk uit te kunnen breiden.

### 3 Casus Innolab

Onderstaande vier opdrachten zijn vanuit twee verschillende thema's geformuleerd; Open data en Fact-Checking. In elk thema worden deze twee thema's op een andere manier verband gebracht.

#### **Onderzoeksthema 1: Open Data**

Den Haag heeft als een van de vier grote steden in de Randstad te maken met zogeheten grootstedelijke problematiek; prachtwijken zoals Duindorp zijn volop en helaas veelal, negatief in het nieuws.

De Haagse wijk Duindorp heeft een slechte reputatie. De Gemeente werkt al jaren aan het verbeteren van het imago van deze prachtige wijk aan zee. Overheden verzamelen al jaren informatie over het leefklimaat van deze wijk, denk aan informatie van de Gemeente Den Haag via Open Data, het CBS en nieuws, maar in hoeverre komt deze informatie nu écht overeen met de negatieve geluiden uit de media en veel gehoorde vooroordelen?

In deze casus onderzoeken de studenten informatie veelal te vinden in Open Datasets die gaan over het leefklimaat van de wijk. En stellen zichzelf vragen binnen het onderzoeksthema (Linked) Open Data.

Hoe levert Open Data een bijdrage aan het verbeteren van het leefklimaat van de Haagse Prachtwijk Duindorp?

#### **Casus 1 Mapping Duindorp**

De studenten brengen in kaart welke overheidsinstanties of initiatieven van de overheid betrokken zijn bij de wijk Duindorp en welke inzichten deze geven over de wijk. Zij hoeven zich niet te beperken tot gegevens van de Gemeente Den Haag, maar kijken ook naar inzichten die iets toevoegen aan de huidige beeldvorming over de wijk.

#### **Casus 2 Kwetsbare straten**

De studenten brengen in kaart welke informatie er beschikbaar is over de wijk Duindorp op straatniveau. Hiervoor kijken ze naar veiligheid, voorzieningen en leefklimaat op straatniveau.

## **Onderzoeksthema 2: Fact-checking**

De HHS werkt samen met Gemeente Den Haag aan de Fact Checking Factory in het tegen gaan van nepnieuws. In deze digitale maatschappij zijn onjuistheden en misinformatie alomtegenwoordig. De HHS heeft met het InnoLab in Duindorp in samenwerking met de Gemeente Den Haag een Fact Checking Factory (FCF) opgezet om verspreiding van nepnieuws tegen te gaan. Nepnieuws wordt onder andere bestreden door het tijdig op te speuren en te ontkrachten. Onderdeel van de FCF is zo het doen van onderzoek naar regionaal nieuws in relatie tot waarheidsbevinding. Welke rol kan ICT spelen in het vaststellen van de feiten?

In dit project onderzoeken de studenten welke rol ICT kan spelen bij het controleren van regionaal nieuws. Aan de hand van online berichten zoals krantenartikelen en sociale mediaberichten onderzoeken de studenten verschillende ICT-technieken en trends.

De centrale vraag voor deze casus is: Hoe kan ICT bijdragen aan het checken van regionaal nepnieuws?

## **Casus 3 mediawijsheid (IMS12)**

De gemiddelde buurtbewoner weet wat er speelt in de wijk, deze informatie is niet altijd online te verifiëren. Aan de andere kant blijken online en sociale mediaberichten niet altijd te kloppen. Hoe kan ICT een bijdrage leveren aan het verbeteren van de mediawijsheid van de burgers uit Duindorp? De 'C' uit ICT kan hier een mogelijke grote rol spelen.

## **Casus 4 Feiten checken van regionaal nieuws**

De Haagse wijk Duindorp heeft een slechte reputatie. De Gemeente werkt al jaren aan het verbeteren van het imago van deze prachtige wijk aan zee. Nu zijn de studenten van de HS aan zet. De studenten werken aan het ontkrachten van regionaal nepnieuws over de wijk Duindorp. Hoe kan ICT, Linked Open Data in het bijzonder, een bijdrage leveren aan het verbeteren van het imago van de wijk?

In de onlinenieuws wereld wordt volop geëxperimenteerd met het toepassen van Linked Open Data. Linked Open Data biedt volop mogelijkheden voor het (her)gebruiken van nieuws. Een veel gebruikte toepassing is het herkennen van namen (named entity recognition), om die vervolgens te identificeren en te linken aan Wikipedia-beschrijvingen. Een voorbeeld van deze technologie is te vinden in het KB Lab in de xportal: <http://kbresearch.nl/xportal>. Waar het mogelijk is om te zoeken op personen, zoals Frans

Timmermans<sup>9</sup> maar ook op alle personen geboren in Maastricht<sup>10</sup>. Het toepassen van Linked Open Data bij het checken van nieuws wordt nog (vrijwel) niet toegepast. Een gemiste kans, door het inhoudelijk samenbrengen van nieuwsberichten kunnen nieuwe verbanden in kaart gebracht worden. Hoe zou een dergelijke toepassing in zijn werk gaan? Of nog beter gezegd in relatie tot Duindorp; welke rol kan Linked Open Data spelen in het controleren van nepnieuws?

## **Bronnen**

- Open Data sets zoals te raadplegen op Open Data platformen.
- Regionale nieuwsberichten en Sociale Mediaberichten (eventueel opvragen via docent: Fenneke Mink)
- Literatuur op het gebied van Big Data, Open Data en Linked Open Data.

---

<sup>9</sup> (<https://www.wikidata.org/wiki/Q32681>, of <https://www.parlement.com/id/vg09llk9rzrp/f c g m frans timmermans>)

<sup>10</sup> (<https://www.wikidata.org/wiki/Q1309>).

## 4 Casus CBS Data voor Wikipedia

Het CBS is uitgegroeid tot een innovatieve organisatie, die actief inspeelt op behoeften van de samenleving. Het deelt haar rijkdom aan informatie met andere overheden, politiek, wetenschap, media en het bedrijfsleven. Zo maakt het CBS het mogelijk dat maatschappelijke debatten en besluitvorming worden gebaseerd op betrouwbare statistische informatie. Het CBS biedt informatie die uitnodigt tot gebruik en zorgt ervoor dat de informatie toegevoegde waarde heeft voor de gebruikers. Niet alleen losse cijfers worden gepubliceerd, maar zoveel mogelijk worden economische en sociale fenomenen beschreven door cijfers in samenhang te presenteren. Het CBS sluit aan bij actuele gebeurtenissen door het uitbrengen van nieuwsberichten en via social media. CBS stelt haar data beschikbaar voor hergebruik aan de hand van verschillende ICT-toepassingen en technieken, zo hanteren zij Open Data standaarden en gebruik van een API.

Artikelen in Wikipedia vormen een belangrijke bron van informatie. Zoekmachines tonen hits uit Wikipedia hoog in de resultatenlijsten. Voor het schrijven van Wikipedia-pagina's wordt om bronvermelding gevraagd. Het CBS geeft net als Wikipedia over veel onderwerpen informatie. Je ziet dan ook vaak dat artikelen mede gebaseerd zijn op informatie van het CBS. Dit kan inhoudelijke informatie betreffen, maar dit kunnen ook referenties zijn naar CBS-artikelen.

Het CBS heeft weinig zicht op het gebruik van CBS content in Wikipedia, denk aan de omvang, de status en de betrouwbaarheid van de referenties. Ook hebben zij geen zicht op welke informatie uit Wikipedia mist in Wikidata. Hoe kan het CBS een bijdrage leveren aan het actualiseren van relevante informatie in Wikipedia artikelen? Denk hierbij aan het linken van Wikipedia artikelen aan entiteiten uit CBS-data. Welke feiten zijn er vastgelegd in Wikidata over deze entiteit? Welke feiten die in Wikidata staan zijn in conflict met de data die het CBS geeft? Wat voor data heeft het CBS die nog niet in Wikipedia is vastgelegd?

Naast het toevoegen van informatie van het CBS op artikelbasis is het mogelijk gegevens toe te voegen aan Wikidata, via de CBS API. Wikidata is de database waaraan Wikipedia artikelen hun informatie kunnen ontleen, de informatie van het CBS zal hier geautomatiseerd aan toe moeten worden gevoegd, hoe zou dit in zijn werk gaan?

De studenten doen tijdens dit onderzoeksproject onderzoek naar de mogelijkheden tot het geautomatiseerd toevoegen van CBS-informatie aan Wikipedia aan de hand de CBS

API. En onderzoeken welke methodes leiden tot de meest zinvolle inzichten voor het CBS en de maatschappij.

### **Casus 1: CBS Content in Wikipedia**

De studenten onderzoeken het Gebruik van CBS content in Wikipedia, de omvang, de status (up to date?) en de aard van de referenties.

### **Casus 2: Automatiseren van CBS content**

De studenten onderzoeken de mogelijkheid om CBS content aan Wikidata geautomatiseerd toe te voegen. De CBS-website heeft een API (volgens het OData protocol) waarmee tekst en data geautomatiseerd kan worden bevraagd. Zie de link naar open data op de website. Volgens Wikimedia is automatische toevoeging (via een API) mogelijk<sup>11</sup>. De studenten onderzoeken voor welke pagina's dit zinvol is en hoe het automatisch toevoegen van content in zijn werk zou kunnen gaan.

---

<sup>11</sup> <https://www.mediawiki.org/wiki/Wikibase/API>

## 5 Casus Open Data voor de lokale overheid

Opdrachtgever Gemeente Den Haag en Provincie Zuid – Holland

De overheid bezit veel informatie. Bijvoorbeeld over wat zichtbaar is in de openbare ruimte, geografische data, statistieken over de bevolking en over toekomstige ontwikkelingen. Een deel hiervan is te gebruiken voor open data. Nieuw aan open data is dat de informatie zo wordt gepresenteerd dat gebruikers deze meteen en vrij, in een 'open standaard', kunnen toepassen. Open standaard betekent dat de informatie voor iedereen te gebruiken is zonder dat daar een apart programma voor nodig is. Hierdoor kunnen ondernemers en anderen gegevens combineren en nieuwe toepassingen ontwikkelen. Dit biedt zakelijke kansen.

Open Data als structureel onderdeel van het primair proces van de gemeente behandelt de volgende thematiek:

- een bestuurlijke en commerciële kijk op het belang van open data;
- het strategisch openen van data;
- de techniek achter open data;
- ethische en juridische aspecten van openheid van data;
- een business case en toekomstige ontwikkelingen van open data.

Onderzoeksthema's voor de Gemeente Den - Haag:

1. Hoe kan Open Data een bijdrage leveren aan het thema leefbaarheid, veiligheid en wonen in Den Haag?
2. Hoe kan Open Data een bijdrage leveren aan het thema verkeer en verkeersveiligheid?
3. Hoe kan Open Data een bijdrage leveren aan sociale voorzieningen en openbare ruimte in Den Haag?
4. Hoe kan Open Data een bijdrage leveren aan het handels- en vestigingsklimaat van de stad?

Onderzoeksthema's voor Provincie Zuid - Holland:

1. Hoe kan Open Data een bijdrage leveren aan het thema Sport & Recreatie voor de provincie Zuid-Holland?
2. Hoe kan Open Data een bijdrage leveren binnen het thema Duurzaamheid & Milieu

3. Hoe kan Open Data een bijdrage leveren binnen het thema Bouwen & Wonen?
4. Hoe kan Open Data een bijdrage leveren binnen het thema Verkeer & Vervoer?

#### Opdracht

Bovenstaande thema's zijn aanleiding voor het onderzoek van de studenten. Zij zullen zelf een hoofdvraag, deelvraag, onderzoeksmethode en datasets kiezen bij het gekozen onderzoeksthema.

Een vraag als "Hoe kan Open Data een bijdrage leveren aan het thema leefbaarheid, veiligheid en wonen in Den Haag?" Kan bijvoorbeeld worden opgesplitst in verschillende deelvragen, zoals:

- Welke factoren dragen bij aan leefbaarheid?
- Hoe meet je leefbaarheid?
- Wat verstaat de gemeente Den Haag onder leefbaarheid?
- Welke gegevens zijn beschikbaar in de data set?
- Etc, etc..



## 6 Casus Beeld en Geluid

### Inleiding

Beeld en Geluid is als instituut voor mediacultuur continu op zoek naar manieren om haar collectie zo breed mogelijk beschikbaar te stellen. Daarnaast probeert het instituut ook om haar collectie en hedendaagse media te duiden door extra informatie aan te bieden aan haar eindgebruikers. Daarvoor gebruikt Beeld en Geluid op dit moment diverse platforms:

- De website: <http://verhalen.beeldengeluid.nl>: hier biedt Beeld en Geluid verdieping rondom historische, maar ook actuele onderwerpen uit de media.
- De website: <https://wiki.beeldengeluid.nl/>: Deze wiki ontsluit kennis over de collectie van Beeld en Geluid en de audiovisuele mediageschiedenis van Nederland. De artikelen worden geschreven door medewerkers van Beeld en Geluid en diverse experts.
- De website: <http://in.beeldengeluid.nl/> is het startpunt voor het verkennen van het archief van het Nederlands Instituut voor Beeld en Geluid. Waar auteursrechten het toestaan is materiaal ook online te bekijken. Op de zogenaamde kanalen wordt content uit het archief thematisch gepresenteerd. Op al deze platforms komen termen voor die ook voorkomen in de thesaurus (de zgn. GTAA) die Beeld en Geluid gebruikt om haar collectie te beschrijven. Deze thesaurus bevat meerdere 'concept schemes': Persoonsnamen, Genres, Onderwerpen e.d. Een deel van de Persoonsnamen is inmiddels gekoppeld aan Wikidata, een open kennis database met gestructureerde informatie die door een gemeenschap wordt onderhouden. Door deze links heeft Beeld en Geluid nu beschikking over de data uit Wikidata waarmee allerlei nieuwe toepassingen mogelijk zijn.

### Opdracht

Verdiep je in één of meerdere van bovenstaande platforms. Voer bijvoorbeeld een SWOT-analyse uit. Op basis van jullie analyse bepalen jullie hoe een Linked Data benadering van waarde kan zijn voor dat platform. Specifiek kiezen jullie voor een doel waarmee de data uit Wikidata ingezet kan worden ter verrijking en versterking van de genoemde platforms. Vervolgens werken jullie dit doel verder uit aan de hand van een degelijke onderbouwing. Jullie bevindingen worden vertaald naar functionele aanpassingen en/of aanpassingen in het redactieproces van het platform. De resultaten worden, waar mogelijk in een (werkende) demonstrator gepresenteerd.

Let op: Voor deze opdracht komen de studenten in week 3 één keer naar Beeld en Geluid in Hilversum om beter zicht te krijgen op de organisatie en de usecase(s) in kwestie. Verder werken de studenten zelfstandig aan de opdracht.

#### Over Beeld en Geluid

In het kenmerkende, veelkleurige gebouw bij de ingang van het Mediapark in Hilversum vind je het Nederlands Instituut voor Beeld en Geluid. Deze cultuurhistorische instelling verzamelt en conserveert audiovisueel erfgoed van nationaal belang. Om deze audiovisuele bronnen toegankelijk te maken voor zoveel mogelijk gebruikers, doet de afdeling Kennis en Innovatie onderzoek naar manieren om de collectie van Beeld en Geluid sneller te ontsluiten en op innovatieve manieren toegankelijk te maken. Zowel mediaprofessionals, onderwijzers en wetenschappers als het algemene publiek moeten immers op optimale wijze toegang kunnen krijgen tot dit historisch en cultureel waardevolle materiaal.

Beeld en Geluid is één van de grootste audiovisuele archieven van Europa. Het instituut beheert ruim 70% van het Nederlandse audiovisuele erfgoed. De collectie bevat meer dan 750.000 uur aan televisie, radio, muziek en filmmateriaal vanaf de vroege twintigste eeuw tot vandaag. Dagelijks stromen alle uitzendingen van de publieke omroep digitaal binnen en ook particulieren en instellingen vertrouwen hun collecties toe aan Beeld en Geluid. Het instituut zorgt ervoor dat het materiaal in een optimale staat blijft voor (her)gebruik.

Beeld en Geluid is een knooppunt in het nationale en Europese audiovisuele domein, een domein dat aan razendsnelle veranderingen onderhevig is - zowel op technologisch als op maatschappelijk vlak. Door de zich snel ontwikkelende informatiemaatschappij veranderen ook de vragen en behoeften van verschillende typen gebruikers van het audiovisuele archief en het museum. De afdeling K&I speelt bij het vergaren van kennis en expertise hierover een belangrijke rol: ze volgt relevante ontwikkelingen op de voet door te participeren in nationale en internationale onderzoeksprojecten, een netwerk van relevante experts op te bouwen en ervaring op te doen met nieuwe technologieën.

## 7 Data Driven Society: mogelijke bronnen

Hieronder treffen de studenten een overzicht aan dataplatforms. Deze lijst is niet uitputtend maar dient als eerste aanleiding bij het zoeken naar data.

### Aanbevolen literatuur:

Onderstaande bronnen zijn aanbevolen bij het Data Driven Society track en dienen als startpunt voor het zoeken van relevante literatuur bij de onderzoeken van de studenten.

- Becker, Peter, et all. "Maak het vindbaar; op schijven, sites en Sharepoint". 2017.  
<https://udocstore.nl/docs/9789492388001>.
- Berners-Lee, Tim. The Semantic Web. Ted 2009.  
[https://www.ted.com/talks/tim\\_berners\\_lee\\_on\\_the\\_next\\_web?language=nl](https://www.ted.com/talks/tim_berners_lee_on_the_next_web?language=nl)
- Bizer, Christian, Heath, Tom and Berners-Lee, Tim (2009) Linked Data - the story so far. *International Journal on Semantic Web and Information Systems*, 5 (3), 1-22. <https://www.igi-global.com/gateway/article/37496>
- Blaney, Jonathan. "Introduction to the Principles of Linked Open Data".  
<https://programminghistorian.org/lessons/intro-to-linked-data>.
- Platform Linked Data Nederland. "Nederlandse Parels van Linked Data Toepassingen". 2015 <http://www.pilod.nl/w/images/e/ee/LinkedDataParels.pdf>.

### Platforms:

- Archeologisch Depot: <https://data.overheid.nl/data/dataset/archeologische-objecten>
- Den Haag Smart City App Store: <http://denhaag.smartcityapp.nl/>
- Data.overheid.nl <http://data.overheid.nl/>
- Dataplatform.nl: <https://denhaag.dataplatform.nl/>
- Den Haag: <http://store.thehagueopendata.org/>
- Geoportaal: <http://geoportaal.ddh.opendata.arcgis.com/>
- Hackalod culturele Linked Open Data: <http://hackalod.com/index.php/data/>
- KB: <http://kbresearch.nl/xportal>
- Linked Data Learning Environment: <http://linkdale.org/>
- Datavisualisatie platform <http://www.informationisbeautiful.net/data/>
- Open Data Protocol: <http://www.odata.org/>
- Data Archiving and Networking Services: <https://dans.knaw.nl/nl/zoeken>

- Data Platform Linked Open Geodata: <https://data.labs.pdok.nl/>
- Data Platform Den Haag: <https://denhaag.dataplatform.nl/data>
- Platform voor ruimtelijke data: <https://inspire.ec.europa.eu/>
- Data Platform CBS:  
[https://opendata.cbs.nl/dataportaal/portal.html?\\_la=nl&\\_catalog=CBS](https://opendata.cbs.nl/dataportaal/portal.html?_la=nl&_catalog=CBS)
- Data platform Civity: <https://www.dataplatform.nl/>
- Data platform Kadaster: <https://www.kadaster.nl/dataplatform> voor de toekomst
- Local Focus data tool: <https://www.localfocus.nl/nl/>
- Open Spending NL platform: <http://openspending.nl/>
- Open vaarroutes: <http://www.varenzh.nl/>
- Openbaar subsidieregister: <https://www.zuidholland.nl/loket/subsidies/subsidies/subsidies/openbaar/>
- Platform Linked Data Nederland:  
[http://www.pilod.nl/wiki/Platform\\_Linked\\_Data\\_Nederland](http://www.pilod.nl/wiki/Platform_Linked_Data_Nederland)
- Provinciaal Archeologisch Depot: <http://www.archeologie.zuidholland.nl/>
- Provinciaal Georegister: <http://www.provinciaalgeoregister.nl>
- Provincie Zuid-Holland  
[http://www.zuidholland.nl/content\\_provincie/content\\_feiten\\_en\\_cijfers/c\\_open\\_data.htm](http://www.zuidholland.nl/content_provincie/content_feiten_en_cijfers/c_open_data.htm)
- Staat van Zuid-Holland: <http://staatvan.zuidholland.nl/Paginas/default.aspx>
- Zoekmachine voor open data in Nederland: <http://opendatanederland.org/>

## 8 Data Driven Society: mogelijke deelvragen

Hieronder treffen de studenten een overzicht van verschillende onderwerpen die hen op weg kunnen helpen bij het formuleren van onderzoeksvragen en deelvragen en nadere onderwerpsbepaling bij het zoeken van literatuur. Deze lijst is niet uitputtend maar is een handig hulpmiddel bij het bepalen van richting tijdens het opstellen van een onderzoeksplan.

Daarnaast dient onderstaande lijst als gesprekstof tussen de studenten, opdrachtgever en docenten tijdens de feedbackmomenten om gezamenlijk tot concrete afspraken te komen richting een wenselijk eindresultaat.

### Overheid open Data

- Onderzoek naar toepassing van Open Data voor opdrachtgever
- Open Data toepassing selecteren aansluitend op de onderzoeksvraag en situatie van de opdrachtgever.
- Ethische en juridische aspecten van openheid van data.
- De toekomstige ontwikkelingen van open data.
- Open Data infrastructuur, tooling en datastores
- Strategisch openen van data: random alles openen of slim aanpakken.
- Waarom Open Data: een bestuurlijke en commerciële kijk op het belang.

### (Community) platform

- (Community) platforms
- Burgerparticipatie
- Doelgroepanalyse
- Informatiebehoefte van de doelgroep

### Semantic web

- Belangrijkste ontwikkelingen op het gebied van het semantische web specifiek binnen bedrijven en overheid.
- De belangrijkste toepassingen van semantische webtechnologieën binnen de overheid.

- Kennis op het gebied van Linked Data (RDF, XML, SKOS, OWL en SPARQL)
- Kennis op het gebied van KOS en Graph searching.

#### Geo-informatie

- Binnen de overheid is meer dan 80% van de informatie aan de locatie gekoppeld en daarmee geo-informatie.
- Geo-informatiemanagement: Welke gereedschappen en welke aanpak moet de informatiespecialist gebruiken om zich geo-informatiemanager te noemen. De voorbeelden laten het vakgebied leven.
- Een geospecialist demonstreert de mogelijkheden van open data aan de hand van GIS en WebGIS oplossingen en de kracht van een goede infographic.
- Datagestuurd werken binnen de overheid en hoe ruimtelijke data ingezet kan worden binnen de slimme stad.
- Overzicht van de ontwikkelingen binnen het vakgebied over een periode van 30 jaar.

#### Verandermanagement ten opzichte van Open Data in overheid

- Theorieën op het vlak van organisatieverandering
- Leiderschapstijlen en hun relatie met verandering
- Verschil en de relatie tussen verandernoodzaak, -geschiedenis en –dynamiek.
- Belangrijke aspecten bij het veranderen van individueel gedrag
- Het duiden van weerstand en relateren aan motivatie, emotie, gezondheid en welzijn bij verandering.

#### Linked Open Data

- Het maken van zinvolle verrijkingen en toepassingen van content.
- Het stimuleren van interactie tussen publiek en cultureel erfgoed.
- Het leveren van een bijdragen aan het Fact-checking van lokaal nieuws.
- Het verbeteren van de kwaliteit van Wikipedia.
- Het inzetten van LOD door culturele instellingen zoals Beeld & Geluid.