

BZK-MVO Open Data Hergebruikers in Beeld 2019

Frederika Welle Donker

f.m.welledonker@tudelft.nl



KENNISCENTRUM
OPEN DATA

<http://www.kcopendata.eu>



Kenniscentrum Open Data (nog wel)

Onderzoek met een focus op governance van open data, impact van open data, juridische en financiële randvoorwaarden voor implementeren van open data beleid

- Open data businessmodellen
- Open (geo)data infrastructuren/ ecosystemen
- Open data (her)gebruik(ers)
- Urban Data Science



KENNISCENTRUM
OPEN DATA

<http://www.kcopendata.eu>



Open Data Boek

<https://link.springer.com/book/10.1007/978-94-6265-261-3#toc>



Overzicht

1. Achtergrond Onderzoek
2. Onderzoeksmethodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

Onderzoek i.o.v. Ministerie van BZK

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

- NL DIGITAAL Data Agenda Overheid*
 - Burgers en ondernemers centraal stellen
 - Gebruikers gedreven overheidsdata
- Maar wie zijn open data hergebruikers?
 - Welke categorieën hergebruikers?
 - Op welke manier worden open data hergebruikt?
 - Waar lopen hergebruikers tegenaan?
 - Welke behoeften hebben hergebruikers?

Onderzoeksmethodes

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

- Kwantitatieve analyse van IP-nummers logfile data januari en februari 2019
 - CBS Open Data Statline
 - PDOK
 - data.overheid.nl
- Sentimentanalyse Tweets tussen 2010 en 2018
- Kwalitatieve analyse
 - Online enquête (23 nov. 2018 – 31 jan. 2019)
 - Workshop wetenschappelijke gebruikers (8 mrt '19)
 - Diepte-interviews met hergebruikers (juni '19)
 - Online validatie van uitkomsten (juli '19)

IP-adres logfile analyses

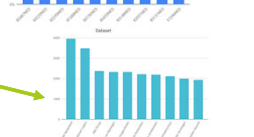
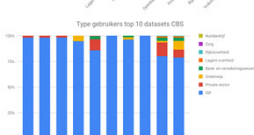
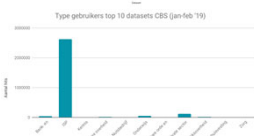
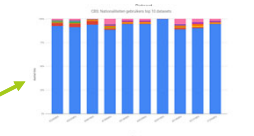
Per platform Top 10 datasets

- Uit welke landen komen meeste requests?
- Welke gebruikersgroepen kunnen we identificeren?
- Welke grootgebruikers kunnen we identificeren?
- Zijn er kruisverbanden?

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

Resultaten CBS Open Data

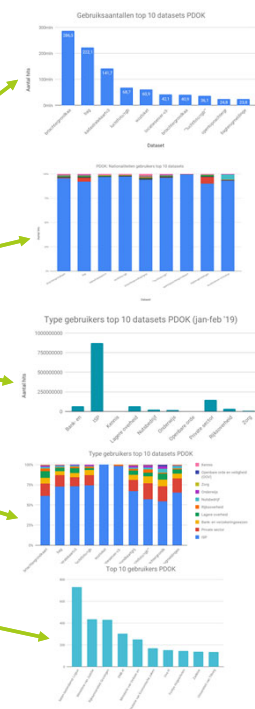
- Kerncijfers Wijken & Buurten, en Bevolking & Huishoudens meest populair
- Relatief veel requests uit buitenland, daarna Duitsland en Hong Kong
- Merendeel gebruikers ISP categorie
- Vooral Consumentenprijzen en kerncijfers populair bij private sector
- Top 10 grootgebruikers vooral publieke sector, hoger onderwijs en bank-/verzekeringswezen



1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

Resultaten PDOK

- Basisregistraties, luchtfoto's en Wozloket meest populair
- Merendeel requests uit Nederland
- Merendeel gebruikers ISP categorie
- Top 10 datasets door allerlei categorieën hergebruikers
- Top 10 grootgebruikers vooral publieke sector, hoger onderwijs en De Nederlandse Bank



1. Achtergrond Onderzoek


2. Onderzoeks-methodologieën


3. Uitkomsten kwantitatieve logfile analyse

4. Uitkomsten kwalitatieve analyse

5. vervolgstappen

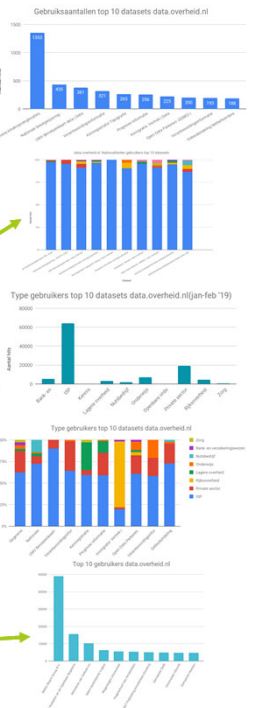
KENNISCENTRUM OPEN DATA <http://www.kcopendata.eu>





Resultaten data.overheid.nl

- Gegevens Kinderopvanglocaties en Verantwoordingsinformatie Woningcorporaties meest populair
- Merendeel requests uit NL, daarna VK en VS
- Merendeel gebruikers ISP categorie
- Redelijke spreiding datasets per categorie
- Top 10 grootgebruikers vnl. overheid, hoger onderwijs en private sector



1. Achtergrond Onderzoek

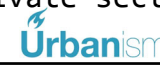
2. Onderzoeks-methodologieën


3. Uitkomsten kwantitatieve logfile analyse

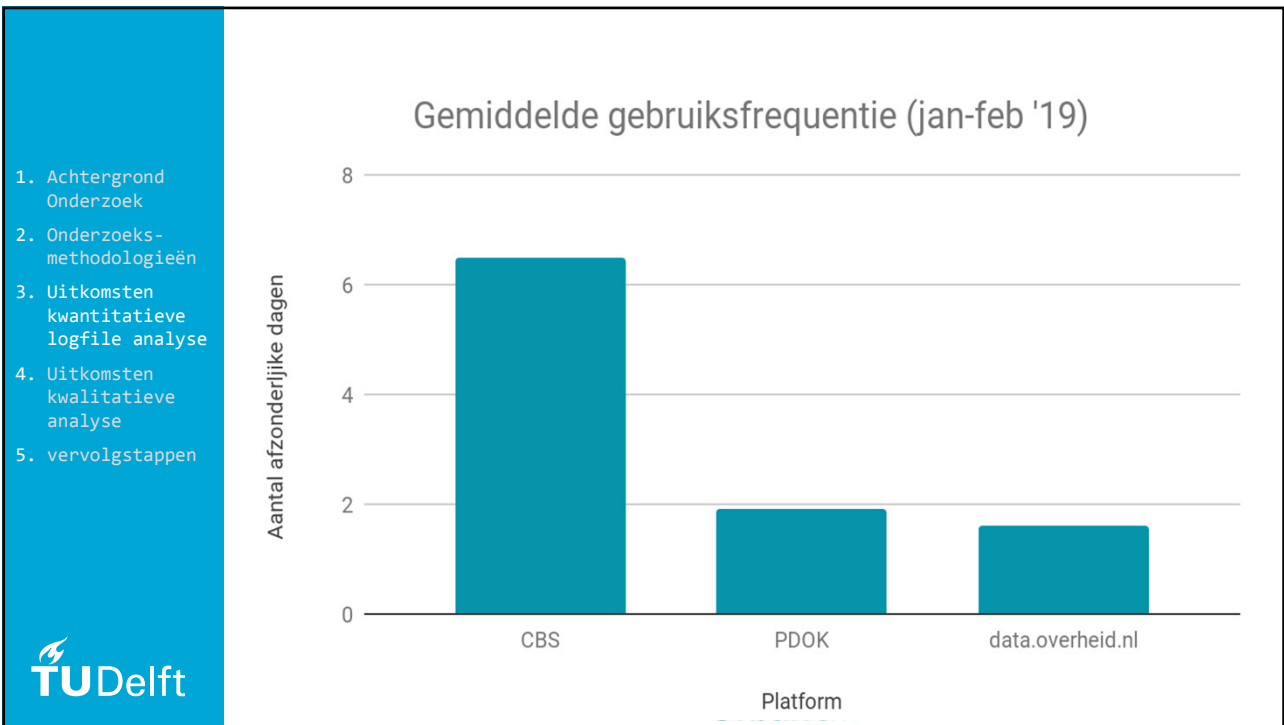
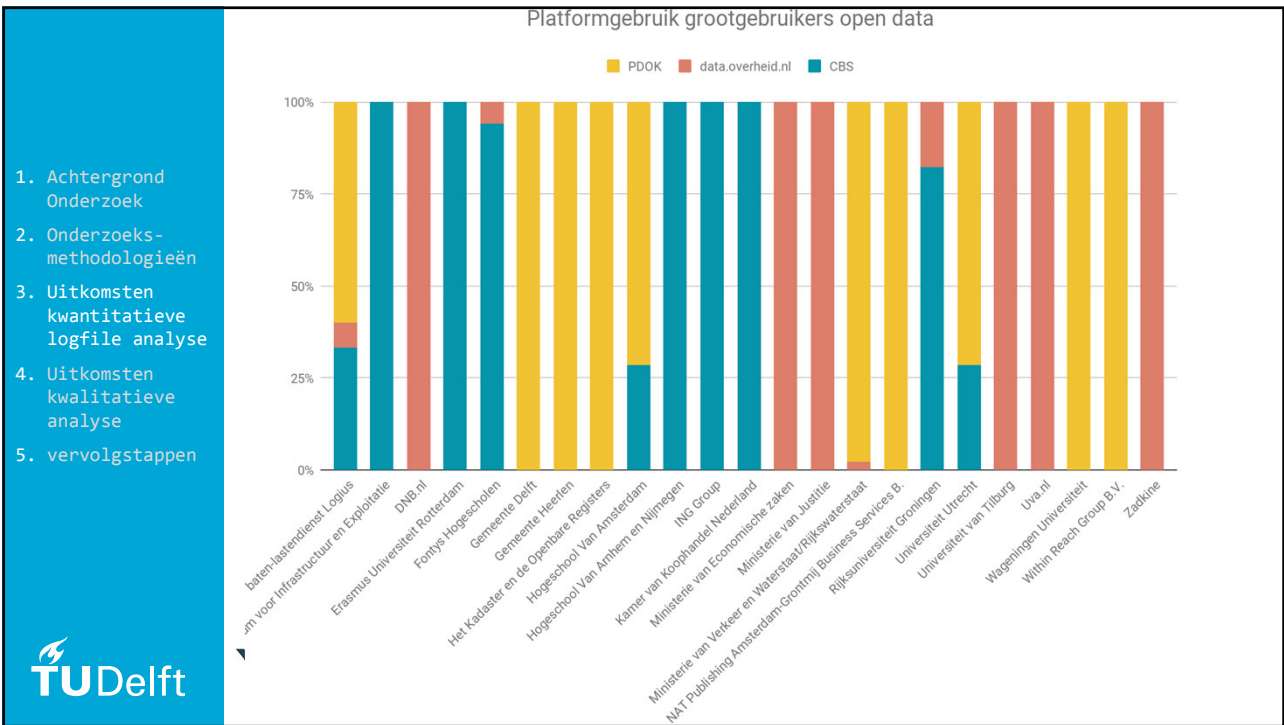
4. Uitkomsten kwalitatieve analyse

5. vervolgstappen

KENNISCENTRUM OPEN DATA <http://www.kcopendata.eu>







1. Achtergrond Onderzoek
 2. Onderzoeks-methodologieën
 3. Uitkomsten kwantitatieve logfile analyse
 4. Uitkomsten kwalitatieve analyse
 5. vervolgstappen
- TU Delft**

1. Achtergrond Onderzoek
 2. Onderzoeks-methodologieën
 3. Uitkomsten kwantitatieve logfile analyse
 4. Uitkomsten kwalitatieve analyse
 5. vervolgstappen
- TU Delft**

Conclusie Logfile Analyse

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

- Meeste gebruikers maken van meerdere platforms gebruik
- Context en samenhang nodig bij logfile analyses
- Ook hergebruikers uit het buitenland
- PDOK data wordt vaker commercieel gebruikt dan CBS data
- Circa 2x zoveel PDOK gebruikers als voor CBS (Statline)
- Gebruik data.overheid.nl vele malen lager dan CBS en PDOK, en vooral als catalogusdienst



Conclusies Twitter sentimentanalyse

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

- Merendeel van Tweets neutraal sentiment
- Twitter wordt vnl. gebruikt voor melden van nieuws en (niet meer) over data
- Meeste commentaar betrof
 - Vindbaarheid
 - Datakwaliteit
 - Beschikbaarheid
- Aantallen te klein om harde conclusies te trekken



<http://www.kcopendata.eu>



Interviews met hergebruikers

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

- 20 hergebruikers geselecteerd na analyse nieuwsberichten over open data gebruik
- 12 hergebruikers geïnterviewd
 - ZZP / micro
 - MKB / grote bedrijven
 - Maatschappelijke organisaties
 - Bijklussende burgers / wetenschappers
 - Kennisinstituut (overheid)

Uitkomsten interviews

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

- Hergebruik voor:
 - Commerciële doeleinden
 - Betaaldiensten / platforms
 - Gratis diensten / platforms
 - Maatschappelijke doeleinden
 - Om de burger te bedienen
 - Als spin-off van andere activiteiten
 - In opdracht van de overheid
 - Om mee te experimenteren

Data zoekstrategieën en vindplaatsen

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

- Zoeken via
 - Google
 - Eigen netwerk
 - Social Media
 - Data catalogi (data.overheid.nl, NGR)
- Vindplaatsen
 - Vooral direct bij de bronhouder
 - Via thematische platforms (CBS, PDOK, NDW)
 - Soms via commerciële bedrijven (Esri)

Meest genoemde belemmeringen

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

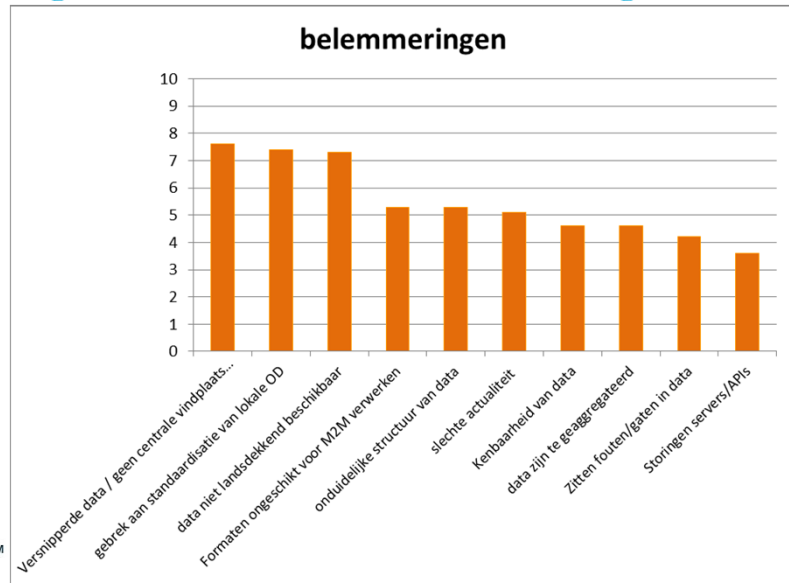
- Vindbaarheid
- Structuur
- M2M verwerkbaarheid
- Actualiteit
- Detailniveau: te geaggregeerd
- Versnipperd aanbod, vooral van lokale overheden
- Bestandsformaten (complexiteit, grootte)
- Summiere metadata / documentatie
- Bestendigheid / lange-termijn beschikbaarheid



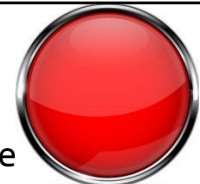
Niks nieuws
dus

Top 10 genoemde belemmeringen

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen



Behoeften van hergebruikers



1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen



- Betere feedbackmogelijkheden en tijdige reactie
- Meer regie op open data van lokale overheden
- Meerdere bestandsformaten / Meerdere diensten
- Meer standaardisatie
 - Datastructuur
 - Internationale standaarden voor API's
- Persistente identifiers
- Data governance
 - Governance van open data
 - CIO/CDO binnen de organisatie)

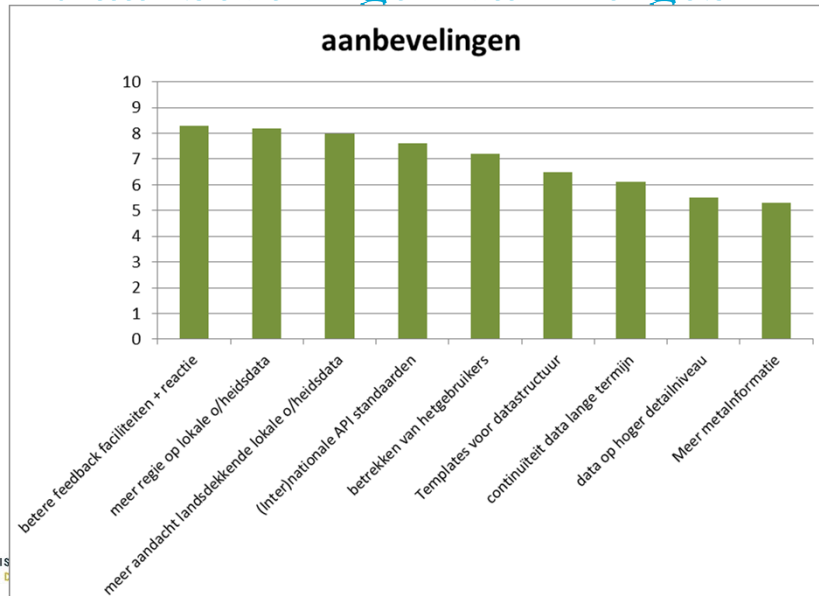
<http://www.kc.opendata.eu>



Top 10 aanbevelingen van hergebruikers

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

TU Delft



Suggesties van hergebruikers

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

TU Delft



- Match actualisatiedata met deadlines voor bedrijfsleven
- Hoger detailniveau voor data via API's (met scraping beveiliging)
- Templates voor gestandaardiseerde en consistente codering voor gelijksoortige datasets
- Voorbeeld van dataset tonen
- Maak meer gebruik van expertise van private sector/ kennisinstituten
- Zet eens vaker de pet van de hergebruik op! Ga bij ze langs!

Betere data governance



<http://www.kcopendata.eu>

Urbanism

Voornaamste conclusies

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

- Verbeter vindbaarheid van data
- Zet in op regie lokale open data governance
- Denk in maatschappelijke vraagstukken en welke datasets daarbij kunnen bijdragen
- Oftewel: verleid die onbekende hergebruiker!

Vervolgstappen

1. Achtergrond Onderzoek
2. Onderzoeks-methodologieën
3. Uitkomsten kwantitatieve logfile analyse
4. Uitkomsten kwalitatieve analyse
5. vervolgstappen

- Logfile analyse aanscherpen
- Tools ontwikkelen om systematische analyses uit te voeren
- Analyse ook voor andere opendataplatforms
- Grootgebruikers interviewen
- Kruisverbanden tussen dataplatforms, datasets en hergebruikers analyseren
- 2e orde effecten van open data onderzoeken

Vragen / suggesties?

f.m.welledonker@tudelft.nl

