Earth, Life & Social Sciences
Van Mourik Broekmanweg 6
2628 XE Delft
P.O. Box 49
2600 AA Delft
The Netherlands

www.tno.nl

T +31 88 866 30 00
F +31 88 866 30 10

TNO report

TNO2014 R10951

Data and the City

Date 27 June 2014

Author(s) Anne Fleur van Veenstra
Jop Esmeijer
Tom Bakker
Bas Kotterink

Copy no
No. of copies
Number of pages 63 (incl. appendices)
Number of appendices 5
Sponsor Ministerie van Binnenlandse Zaken en Koninkrijkrelaties
Project name Scoping study Data Markets
Project number 060.01763

Samenvatting

In 2050 woont 70% van de wereldbevolking in steden.[1] Stedelijke gebieden zullen daarom nog belangrijker worden als centra voor economische ontwikkeling, kennis en creativiteit, maar ze worden ook vaak gekenmerkt als gebieden waarin sociale tweedeling plaatsvindt, waar armoede is als gevolg van werkloosheid en waar een grote aanslag wordt gepleegd op het milieu en de leefomgeving. Digitalisering en informatisering van de harde en zachte infrastructuur leveren veel data op over de fysieke omgeving, diensten, en interacties. Daarom proberen steden, onder de noemer van *smart cities*, steeds vaker te innoveren op basis van deze data door ze te integreren in de infrastructuur en door een 'systeem van systemen' te creëren. Tegelijkertijd hebben beleidsmakers, bedrijven en burgers concrete vragen die beantwoord kunnen worden door een combinatie van data, bijvoorbeeld door informatie op een kaart te plotten, waardoor een slimme informatielaag wordt toegevoegd aan de stad. Er zijn echter niet alleen voordelen; zo kan bijvoorbeeld de privacy van mensen hierdoor in het geding komen.

Ontstaan van een datalandschap
Hoewel deze studie niet primair over smart city ontwikkelingen gaat, richt deze studie zich wel op innovatie rondom data in de stad. Data is door sommigen het "nieuwe goud" genoemd.[2] 'Data-gedreven innovatie' is ontstaan rondom drie trends: big data, open data en het Internet of Things (IoT). Waar big data vooral over nog sterkere processing technologie voor steeds complexere datasets gaat en open data een sterk publiek gedreven trend is, gaat het IoT over de integratie van de fysieke en de digitale wereld, bijvoorbeeld door het toenemende gebruik van sensoren. Door het toenemende gebruik van data bij het nemen van beslissingen, is het van belang inzicht te krijgen in het 'datalandschap' waarin deze beslissingen gemaakt worden, zowel door in kaart te brengen welke data wordt gebruikt en voor welk doeleinden, maar ook door te kijken naar wie er met wie samenwerkt en hoe dit publieke waarden als democratie en autonomie beïnvloedt. Vanwege de breedte van het onderwerp en de behoefte aan inkadering, richt deze studie zich op het verkennen van het datalandschap dat ontstaat op het gebied van leefbaarheid in Rotterdam: de fysieke leefomgeving, veiligheid en sociale cohesie.

Dit onderzoek maakt gebruik van Actor-Netwerk Theorie (ANT). ANT komt voort uit de sociologie en richt zich op het in kaart brengen van netwerken van actoren en hun interacties. Technologieën of artefacten kunnen daarbij ook als actor gezien worden omdat ook zij het netwerk en de uitkomsten kunnen beïnvloeden. Het onderzoek betreft dus een verkenning die zich richt op het – bottom-up – in kaart brengen van ontwikkelingen en bestaat uit drie stappen. Als eerste zijn 33 partijen geïnterviewd (zie Annex 3). Vervolgens worden analyses uitgevoerd naar de relatie van actoren met leefbaarheid, de manier waarop zij waarde creëren met data en de onderliggende toepassingen en strategieën met betrekking tot data en hun ambities en uitdagingen. Op basis hiervan is het data ecosysteem rondom leefbaarheid in kaart gebracht. Deze studie maakt onderscheid tussen het ontstaan van een

[1] Europe in a changing world – inclusive, innovative and reflective Societies, HORIZON 2020 WORK PROGRAMME 2014 – 2015, European Commission Decision C (2014)2690 of 29 April 2014
[2] Kroes, N. (2013). Data is the new gold. Opening Remarks, Press Conference on Open Data Strategy, http://europa.eu/rapid/press-release_SPEECH-11-872_en.htm

datalandschap waarin ontwikkelingen op het gebied van data zich voordoen, en het data ecosysteem, wat een specifiek systeem van partijen beschrijft die samenwerken en daarbij data, kennis en technologieën delen ten aanzien van een specifiek doel, zoals leefbaarheid in de stad. Ten slotte wordt gekeken welke impact de verwachte 'dataficatie' van het landschap heeft op het domein van leefbaarheid in de stad.

Datatoepassingen en -strategieën, ambities en uitdagingen
Het huidige data landschap wordt gekenmerkt door drie typen partijen die een rol hebben die betrekking heeft op leefbaarheid in de stad: publieke partijen (zoals de gemeente, gemeentelijke diensten, de politie en het Rijk), die ook initiatieven financieren (Buurt Bestuurt project), semipublieke partijen als (zorg)verzekeraars en woningbouwcorporaties, en private partijen en dienstenaanbieders (zoals Funda, Iens, Marktplaats en Peerby). De eerste twee groepen zijn primair verantwoordelijk voor leefbaarheid, vaak doordat ze vanuit hun (semi-)publieke taak zijn aangewezen om een aspect van leefbaarheid te realiseren of bewaken. Zo is de politie verantwoordelijk voor de veiligheid in de stad, woningbouwcorporaties voor het realiseren van betaalbare huisvesting en Stadsbeheer, een gemeentelijke dienst voor de buitenruimte en het verwijderen van afval. De derde groep partijen heeft geen primaire taak ten aanzien van de leefbaarheid in de stad, maar hebben soms wel veel data over de stad die mogelijk waardevol zouden kunnen zijn voor partijen die primair verantwoordelijk zijn voor leefbaarheid. Daarnaast leveren ze diensten die een stad aantrekkelijker voor de bewoners maken.

De meerderheid van de partijen die is geïnterviewd voor deze studie gebruikt data voor het primaire proces van de organisatie: ofwel voor de dagelijkse uitvoering van hun taak, ofwel voor tactische en soms strategische doeleinden (bijvoorbeeld voor managementinformatie). De wijze waarop data door deze partijen wordt gebruikt hangt dan ook sterk samen met hun 'business case'. De stap voor partijen om echt nieuwe diensten op basis van data te ontwikkelen is vaak nog groot. Een uitzondering is TomTom, die de verandering heeft ondergaan van een organisatie die primair kaartinformatie verkoopt naar een organisatie die zich daarbij specialiseert in het up-to-date houden van kaartinformatie en die deze dienst integreert met informatie van andere partijen die daarmee inzichten verkrijgen die ze eerst nog niet hadden. Met betrekking tot de strategie die partijen hanteren bij het verwerken van data kunnen er twee belangrijke doeleinden worden onderscheiden: het gebruik van data teneinde de operatie – en het maken van beslissingen – te automatiseren en het contextualiseren en/of personaliseren van data om deze beter te kunnen begrijpen. De meest gehanteerde manier om data te contextualiseren is door middel van geo-informatie. Partijen die zich vooral hebben gericht op het inrichten van een infrastructuur of platform, in plaats van op het ontwikkelen van nieuwe diensten, lijken het best in staat om deze beide strategieën te hanteren.

De belangrijkste ambities met betrekking tot data zijn: (1) het faciliteren van participatie met bewoners en het integreren van sociale media, (2) contextualisatie van data, bijvoorbeeld door profiling, het doen van voorspellingen en personalisatie, (3) het ontwikkelen van nieuwe producten en diensten (ondanks dat partijen data nu nog vooral voor hun huidige operationele doeleinden inzetten) en (4) de verbinding maken met de fysieke wereld en aan te sluiten bij het Internet of Things. Vooral de ontwikkeling van draagbare technologie als smart phones en slimme horloges lijkt hierop van invloed te zijn. De belangrijkste uitdagingen die zijn

genoemd zijn: (1) privacy en gegevensbescherming, (2) de bewustwording van organisaties rondom de mogelijkheden die datatoepassingen kunnen bieden, (3) het daadwerkelijke toepassen van data en data analytics, wat om zeer geavanceerde kennis en vaardigheden vraagt en (4) datakwaliteit.

Data ecosysteem rondom leefbaarheid in de stad
De contouren van een data ecosysteem op het gebied van leefbaarheid in de stad Rotterdam kunnen nu worden geschetst. In een data ecosysteem werken organisaties uit de publieke, semipublieke en de private sector samen bij het realiseren en ontwikkelen van diensten op basis van data. Het generieke ecosysteem kan worden gezien als een aantal lagen van activiteiten (zoals collectie, opslag, validatie, integratie, verrijking, visualisatie) op basis van data. Bij elke laag wordt er waarde toegevoegd. Partijen kunnen zich voor één of enkele specifieke datasets op alle lagen richten om waarde te genereren (verticale diensten), of partijen kunnen zich richten op een specifieke activiteit met betrekking tot data, zoals het inrichten van een platform of specialisatie in visualisaties (horizontale diensten). Het data ecosysteem rondom leefbaarheid in de stad laat zien dat publieke en semipublieke partijen met een primaire taak op het gebied van leefbaarheid zich vaak richten op een ‘kolom’ van activiteiten rondom één dataset en dat er minder partijen zijn die zich op één specifieke activiteit richten. Daarnaast ontstaan er een aantal secundaire rollen rondom het ecosysteem, bestaande uit toezichthouders, durfkapitaal en accelerators, (lokale) media, en (lokale) samenwerkingsverbanden.

Op het gebied van leefbaarheid in de stad wordt nog weinig samengewerkt op basis van data tussen de publieke en de private sector. Wel ontstaan er initiatieven die laten zien dat partijen in deze sector het belang van samenwerking op het gebied van data en data-integratie steeds meer gaan inzien. Voorbeelden zijn de integratie van de Beter Buiten app in het systeem waarmee Stadsbeheer de buitenruimte schouwt en de koppeling van (publieke) gegevensregistraties met de informatiesystemen van de politie die daar gebruik van maakt voor operationele doeleinden. Behalve voor deze operationele doeleinden, worden gegevensbronnen van verschillende partijen ook geïntegreerd voor het gebruik in monitoringsdoeleinden, bijvoorbeeld bij beleidsevaluaties. Opvallend is dat de rijkdom aan data die in het bezit zijn van publieke organisaties en dienstaanbieders (zoals Funda, Iens en Marktplaats) hun weg nog niet vinden naar het primaire proces van leefbaarheid (de operationele systemen van de (semi-)publieke partijen of naar de beleidsevaluaties of –monitoring).

Veel datatoepassingen zijn gebaseerd op een combinatie van transactionele databases en geo-informatie. Veel partijen zijn op dit moment bezig om hun operationele data locatie-gebonden te maken en zo beter inzicht te krijgen in deze data. Typen data die partijen op termijn willen gaan gebruiken of integreren in hun activiteiten zijn social media data en sensor data, om de contextualisatie van hun eigen data ‘sociaal’ of real-time te maken. De integratie van media content als video’s wordt vaak lastiger gevonden, ook vanwege de grootte van de bestanden en ook het gebruik van persoonsgegevens wordt minder genoemd door de geïnterviewde partijen, vanwege de gevoeligheid. Daarnaast valt op dat er wel interactie is tussen open data vanuit de publieke sector en data die eigendom is van private partijen, maar deze interactie vindt doorgaans plaats doordat private partijen open data hergebruiken in hun diensten of doordat de publieke sector vanwege hun wettelijke taak data krijgt van private partijen die ze nodig hebben om

deze taak uit te voeren. Er zijn dan ook nog geen platformen die de uitwisseling van datasets van private partijen structureel ondersteunen, zoals het Rotterdam Open Data portal dat doet voor gemeentelijke datasets.

Terwijl de data infrastructuur meer en meer een ‘commodity’ wordt, vindt de strijd plaats om de toegang tot de data, het liefst zo lokaal mogelijk. Zo wordt er vaak gebruik gemaakt van (het liefst zeer) lokale gegevens terwijl de technologieën die gebruikt wordt om deze gegevens op te slaan, te analyseren en te verwerken worden vaak wereldwijd beschikbaar zijn. Tegelijkertijd proberen organisaties wel de juiste kennis en expertise op te bouwen wanneer zij de het gebruik van data cruciaal achtten voor hun positie in de markt, om zo niet afhankelijk te zijn van derde partijen om waarde uit hun data te halen. Verder is er een strijd gaande om welke organisaties de ‘interfaces’ van de stad in handen krijgen, vanuit verschillende domeinen. Wie immers het platform in handen heeft waarop anderen weer diensten ontwikkelen (de smartphone, de smart car), heeft toegang tot de data én kan de rol op zich nemen van broker en zo een percentage van de diensten opstrijken. Ook ontwikkelen meerdere partijen authenticatiediensten om zo toegang te krijgen tot gegevens.

Conclusies en aanbevelingen
De impact van data-gedreven innovatie op het domein van leefbaarheid in de stad lijkt op dit moment nog gering: (semi-)publieke organisaties blijven het meest bepalend voor de diensten die worden ontwikkeld en geleverd en data wordt daarbij ingezet om bestaande activiteiten te ondersteunen. Tegelijkertijd is wel goed zichtbaar dat zeer veel interessante data die gebruikt kunnen worden voor diensten op het gebied van leefbaarheid in handen zijn private partijen. Aangezien data-integratie steeds belangrijker wordt voor dienstverlening en nieuwe diensten, lijkt de rol van private partijen in de toekomst groter te worden. Dit wordt mogelijk nog eens versterkt door de toenemende invloed van sociale media en sensordata die vaak ook in handen zijn van private partijen. Er lijkt een verschuiving plaats te vinden naar smart city systemen die het best omschreven kan worden van G2C naar G2B2C naar C2B2T2C2All, waarbij niet alleen sensor data, maar ook data die door consumenten en burgers zelf (actief en passief) worden gegenereerd een rol gaan spelen.

Dit is echter pas een van de eerste empirische studies naar smart cities en het ontstaan van een datalandschap in de stad. Hoewel we de huidige ontwikkelingen in kaart gebracht hebben, is er nu nog weinig zicht op de uiteindelijke impact van deze ontwikkelingen. Er blijft dan ook onderzoek nodig naar welke invloed deze ontwikkelingen zullen hebben op de fysieke wereld.

Contents

1 Introduction

By 2050, 70% of world population – and 86% for OECD countries – will live in urban areas.[3] Cities are very important in policies aiming to create growth, jobs and a sustainable future. They are centres of economic development, services, knowledge and creativity, but they are also the places of social polarisation, intercultural confrontations, poverty concentration, unemployment and environmental problems. The European commission, in its H2020 research programme on inclusive societies, defined as a key challenge the "identification of means and ways to make the city an emblematic place for attracting jobs and economic activities, transforming it into a hub of innovation and ensuring social cohesion and cultural dialogue while preserving natural resources and limiting environmental damage for the next generations".

1.1 Smart cities

Ongoing digitisation of soft and hard city infrastructures increases the availability of a wide range of information about physical environments, services, and interactions between people. Cities, often adopting the banner *Smart Cities,* are increasingly trying to leverage this data to align and integrate infrastructure, planning and management, and human services as a system of systems – with the goal of making cities more desirable, liveable, sustainable, and green. Examples of proposed (or sometimes realised) ideas and solutions are smart, distributed energy grids, predictive policing, self-monitoring sewers and more efficient and adaptive transportation systems. A constant in the various applications and visions is the idea that the analysis of data collected via the web, social networks, mobile phones, CCTV camera's and in numerous sensors in roads, cars, networks and devices provide real-time, actionable insights that enable us to improve the way we live: "For the first time we'll see cities as a whole the way biologists see an organism - instantaneously and in excruciating detail, but also alive. Today we see them the way astronomers see heavenly bodies - as it was, some time ago, light-years in the past."[4]

[3] Europe in a changing world – inclusive, innovative and reflective Societies, HORIZON 2020 WORK PROGRAMME 2014 – 2015, European Commission Decision C (2014)2690 of 29 April 2014

[4] Townsend, A. (203) Smart cities: big data, civic hackers and the quest for a new utopia. W.W. Norton & Company Inc., New York, p. 72

BOX 1: Smart Cities

The EU defines the smart city as "a system of people interacting with and using flows of energy, materials, services and financing to catalyse sustainable economic development, resilience, and high quality of life; these flows and interactions become smart through making strategic use of information and communication infrastructure and services in a process of transparent urban planning and management that is responsive to the social and economic needs of society".[5] In the EU-28 around 90% of cities over 500.000 inhabitants implement smart city programmes. This number drops to 51% for cities with 100.000 inhabitants or more. Becoming smart is a widely shared ambition of bigger cities in the EU and the OECD.[6]

Managing and exploiting massive, heterogeneous datasets generated across city subsystems in an effective and meaningful way is an enormous challenge. Furthermore, data about cities and their citizens are being collected by myriad players with different agendas for all kinds of purposes. Public organisations might collect and use data to gain a better understanding of societal challenges and (the impact of) policy interventions, while private organisations may be more focused on insights regarding their customers or new markets. In its Work Programme, the European Commission emphasised the importance of inclusive and trustworthy digital societies.[7] As data becomes a driving force in decision-making by these actors, it is important to get a better understanding of this data landscape, both in terms of the data that is being used, but also *how* it is being used and how this affects public values such as democracy and autonomy. According to Mark Graham from the Oxford Internet Institute "[…] It is important to understand who produces and reproduces, who has access, and *who* and *where* are represented by information in our contemporary knowledge economy." Furthermore, as data and data analytics gain importance, it should be transparent what assumptions are encoded into algorithms that will guide the actions of city planners, public officials or other kinds of players.[8] [9] [10] [11]

While the main focus of this study is not *smart cities*, it does investigate related trends and developments. We will explore the 'data landscape' of cities in their efforts to maintain and improve of quality of life, specifically the City of Rotterdam. The objective is to map how key players are deploying data and data analytics, the data sources and products and services they use, the barriers they encounter and

[5] European Innovation Partnership on Smart Cities and Communities - Strategic Implementation Plan, 2014, http://ec.europa.eu/eip/smartcities/files/sip_final_en.pdf
[6] Mapping Smart Cities in the EU, European Parliament, DIRECTORATE GENERAL FOR INTERNAL POLICIES, POLICY DEPARTMENT A: ECONOMIC AND SCIENTIFIC POLICY, IP/A/ITRE/ST/2013, 02 January 2014
[7] Europe in a changing world – inclusive, innovative and reflective Societies, HORIZON 2020 WORK PROGRAMME 2014 – 2015, European Commission Decision C (2014)2690 of 29 April 2014
[8] Townsend, A. (2013) Smart cities: big data, civic hackers and the quest for a new utopia. W.W. Norton & Company Inc., New York, p. 297
[9] boyd, d. & Crawford, K. (2012). Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. In: *Information, Communication, & Society* 15:5, pp. 662-679
[10] Wakefield, J. (2012). Can we trust the code that increasingly runs our lives? In: *BBC*, http://www.bbc.co.uk/news/technology-19347122
[11] Richards, N. & King, J. (2013). Three Paradoxes of Big Data. In: *Stanford Law Review Online 41 (2013)*, http://www.stanfordlawreview.org/online/privacy-and-big-data/three-paradoxes-big-data

the changes that take place in the value network. Based on this mapping exercise, we aim to answer questions about the impact of data-driven innovation on a specific sector. To allow for detailed insights we focus in this study on the notion of 'leefbaarheid' (liveability) in the city of Rotterdam.

1.2 Data-driven innovation and the data landscape

Data-driven innovation is spurred by a number of interrelated trends: big data, open data and the emergence of the Internet of Things (IoT), and an increasingly popular belief that the use of data and data analytics holds great potential.[12] [13] Big data developments, characterised by the 'three v's': volume, velocity, and variety,[14] are driven by a growing demand and desire for more and better insights and by the availability of more advanced, and cheaper processing technologies that can deal with the explosion of huge, fast and unstructured datasets in recent years. Open data developments, often driven by the public sector aiming to increase transparency and accountability, create more effective decision making and stimulate re-use for economic purposes.[15] The IoT refers to the use of sensors in many applications in the physical world, such as robots for smart manufacturing, mobile phones for location-based services, and home appliances that let you know when the heating should be turned on or off.

All these data sources lead to an enormous increase of available data that can be combined, processed, visualised and interpreted. For example, data can be connected to their environment and visualized in maps, adding a smart information layer. As insightful as this may be, it may also have undesirable consequences for the privacy of individuals, as it becomes possible to get better insights on where citizens are at a given time. While data-driven innovations are expected to have an economic potential as innovations will lead to new services and to the creation of new jobs, not all consequences are likely to be received positively. Besides mapping the organisations in the data landscape that stems from the aggregate of data-driven innovations, this study will also investigate drivers and barriers of these developments, as well as expected consequences and future developments.

1.3 Leefbaarheid

The focus of this study is 'leefbaarheid'. This notion, often translated as 'liveability', and referring to the quality of life in a certain area, encompasses a wide range of aspects, such as the state of the 'leefomgeving' (the physical environment), social aspects such as social cohesion, public safety, the amenities that the neighbourhood offers, interaction with neighbours and the means of transport such

[12] Kroes, N. (2013). Data is the new gold. Opening Remarks, Press Conference on Open Data Strategy, http://europa.eu/rapid/press-release_SPEECH-11-872_en.htm
[13] Asay, M. (2013). Gartner on Big Data: Everyone's Doing It, No One Knows Why. In: ReadWrite, http://readwrite.com/2013/09/18/gartner-on-big-data-everyonesdoing-it-no-one-knows-why#awesm=~ojbBSXmHZo7DQA
[14] Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. In: *META Group*, 6 February, http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf
[15] Veenstra, A.F. van & Broek, T.A. van den (2013). Opening Moves – Drivers, Enablers and Barriers of Open Data in a Semi-Public Organization. In: Wimmer, M.A., Janssen, M. Scholl, H.J. (Eds.), EGOV 2013, LNCS 8074, pp. 50-61

as the state of the roads and public transport facilities.[16] The notion of liveability can be interpreted on different levels: on a street level, neighbourhood level, on the level of boroughs and on the city level. Organisations gather data on these levels for different purposes. The Ministry of the Interior and Kingdom Relations, for example, gathers data on liveability on a national scale (although it does provide analyses on a neighbourhood level) every two years in the '*Leefbarometer*'.[17] This information is used for policy making. The same goes for organisations in the private sphere, who can use data to make better decisions or to profile and target their customers more effectively.

Data can interact with liveability in a number of ways. Big data developments and services can create better insights by mapping developments or services or by combining different data sources to create deeper insights. Sensor data gathered through mobile phones can show the location of people. Blogs or social media can be used to interact with other citizens based on such location data (e.g. *Foursquare*). Furthermore, patterns and profiles of citizens, or profiles of specific areas in terms of safety can be created on an aggregated level, for example by combining data from transactional databases, social media data, and real-time data from sensors. Especially when new techniques and algorithms enable the combination and analysis of data, such as a combination of data on public safety and the physical environment, this could lead to better insights into how different aspects of liveability are connected. Examples include the housing website *Funda* or the *Leefbarometer*. Websites such as *Tinder, Couch surfing* or *AirBnB* also enable people to get in contact with others.

This study focuses on the local and very local level: the city, the neighbourhood, and the street level. As mentioned above, liveability is a very broad concept that encompasses multiple aspects ranging from safety to affordable housing and from mobility to social cohesion. To allow for detailed insight and for creating an overview that is as complete as possible, we focus on three specific aspects of liveability: the physical environment, public safety and social cohesion. The physical environment concerns the appearance of objects in the public sphere. Public safety concerns the domain of crime and the police, but it also includes the perceived security by citizens. However, anti-terrorism and homeland security, which have a more national or even international focus, are not taken into account. Social cohesion is a less tangible concept and it includes both aspects of care and social welfare, the interaction of citizens with their social environment and their neighbourhood. The choice for these aspects is a pragmatic one to scope our study, rather than a choice regarding which aspects have a stronger impact on liveability.

1.4 Rotterdam

While many cities develop data-driven tools regarding liveability (see, for example, Box 2 on Almere's monitor), this study focuses on the city of Rotterdam. The city is often seen as a living lab for social trends; Rotterdam is known, for example, for its unorthodox, front-runner mentality regarding urban development. With around

[16] Leidelmeijer, K. et al. (2008). De Leefbarometer. Leefbaarheid in Nederlandse wijken en buurten gemeten en vergeleken, http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2008/05/01/rapportage-instrumentontwikkeling.html
[17] http://www.leefbaarometer.nl/

620.000 inhabitants,[18] the city is the second largest city in the Netherlands. The city is run by a board of the mayor and aldermen, who are subject to scrutiny of a city council. Furthermore, the city is divided into fourteen areas that have some autonomy in the areas of the physical environment and social welfare. Home to the largest port of Europe, the city has a very international population (166 nationalities in 2012).[19] 37% of the inhabitants have at least one parent of non-Western origins.[20] Furthermore, the population of Rotterdam is more often unemployed than the national average (58% of the population is employed, compared to 65% nationally).[21]

After the old city centre burned down after an air raid in 1940, the city has literally become an urban laboratory for modern architecture. This makes the city a place where innovative solutions can make a difference regarding the environment. Regarding the use of data, liveability, and especially the three topics that were identified, a number of municipal policies are of specific interest to this study. Firstly, regarding the use of data in the city, the municipality developed, together with the Hogeschool Rotterdam, a university of applied sciences, an open data portal.[22] The municipality is actively involved in offering open data via this portal, which has led to the development of a number of apps, such as the *Bomenspotter* (Tree spotter) app.[23] Regarding public safety, fifteen years ago the city had a reputation for being unsafe. Several initiatives were undertaken to increase (perception of) safety in the city, such as the *Veiligheidsindex* (Safety Index)[24] and the *Buurt bestuurt* initiative ('Neighbourhood Governs').[25] In relation to social cohesion, the municipality is currently undergoing changes, as a number of tasks in this domain, such as care for children who cannot live with their parents, have been shifted from the national government to the local level.

1.5 Research question objectives and outline

The main goal of the study is to explore the data landscape of liveability in the city of Rotterdam. The main research question is: 'What does the data landscape of liveability in the city of Rotterdam look like and does the emergence of a or any data ecosystem have any impact on the domain of liveability in the city?'. Answering this research question is done in five steps. Firstly, the domain of liveability in the city is investigated by mapping the main players in this domain. Secondly, the data landscape in relation to liveability in the city is explored by looking at all actors that have data and that develop or use data products on this domain and by determining their main ambitions and challenges. On the basis of this exploration, the third step is to investigate what the data ecosystem of liveability in the city looks like. Subsequently, the impact of these developments within the data landscape on the domain of liveability is investigated. Finally, conclusions on the development of a

[18] According to CBS-data Rotterdam has 617.693 inhabitants on January 1, 2014, http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=37230ned&D1=0-17&D2=70,92,480&D3=142,148-155&VW=T, published on February 6 2014, viewed February 19 2014
[19] Gemeente Rotterdam, Economische kerngegevens Rotterdam 2012, http://www.rotterdam.nl/economischekerngegevensrotterdam2012
[20] NRC Handelsblad, 'Vriendjespolitiek is er ook bij CDA en VVD', March 4, 2014
[21] NRC Handelsblad, 'Vriendjespolitiek is er ook bij CDA en VVD', March 4, 2014
[22] http://www.rotterdamopendata.nl
[23] http://www.bomenapp.nl.
[24] http://www.rotterdam.nl/veiligheidsindex2014.
[25] http://www.buurtbestuurt.nl/

data landscape and data ecosystem within the domain of liveability in the city (Rotterdam) are drawn.

This study, thus, both uses the term data landscape and data ecosystem. Data landscape refers to the general data-related developments. The concept of an ecological approach to describe business environments was introduced by Moore to describe how companies should not be viewed as members of a single industry "[…] but as part of a business ecosystem that crosses a variety of industries."[26] In these ecosystems, collaborative arrangements of firms combine their individual offerings to create coherent, customer-facing solutions.[27] This seems a suitable perspective to explore the dynamics of networks of human and non-human actors, that have started to form around specific data-driven innovations, and that may gradually link together into an all-encompassing (big) data ecosystem. To contextualize the findings from this study, a number of related trends are mapped and analysed, such as the deployment of data sources.

To allow for an extensive investigation and detailed mapping of developments, we use a combination of a top-down and bottom-up approach, with a focus on the latter. This is reflected in our research methodology, which draws on Actor-Network Theory (ANT).[28] Our research approach is not driven by hypotheses, but it is an investigation of what is happening in the field of data-driven innovation in relation to liveability in the city. We will also look at *impact* of the data-related developments on the domain that we study, but we do not attempt to investigate the *effectiveness* of data-driven innovation. The main focus of this study is thus explorative rather than an attempt is made to answer clear research questions. The research approach and methodology of this study will be explained in more detail in the next chapter.

The remainder of this study is structured into four chapters. In Chapter 2, the methodology is explained in more detail. Chapter 3 presents the analyses of the data landscape based on our empirical research, such as the key actors, the main value propositions of data and data analytics, and the ambitions and challenges of the main actors. Chapter 4 discusses the emergence of a data ecosystem and describes a number of findings based on the analyses in the previous chapter. Chapter 5 presents conclusions on the most important impacts of datafication of the domain on (the organisation of) liveability in Rotterdam and its restructuring effects and looks ahead to future developments and future research.

[26] Moore, F. (1993). Predators and Prey: a new ecology of Competition. Harvard Business Review, May-June, http://blogs.law.harvard.edu/jim/files/2010/04/Predators-and-Prey.pdf
[27] Adner, R. (2006). Match your innovation strategy to your innovation ecosystem. Harvard Business Review, April, http://pds12.egloos.com/pds/200811/07/31/R0604Fp2.pdf
[28] See, for example, Latour, B. (2005). Reassembling the social, Oxford: University Press

BOX 2: Almere's monitor: data-driven innovation in the public sector:

The municipality of Almere develops a new liveability monitor: the 'Straatkubus' ('Street cube'), which groups data from the physical, social and safety domain to allow analyses on street level. The objective of the Straatkubus monitor is to lower the costs of solving liveability related problems by doing analyses on neighbourhoods and signaling issues early on. Besides being an instrument for analyses, it functions as a communication tool for local partners, such as welfare organisations that are active in these neighbourhoods. These local organisations are supported by allowing for better insights into the problems in an area and helps them to identify potential partners for collaboration to solve these liveability issues.

The instrument performs data analyses in order to test hypotheses. Firstly, these are used engage in discussions on specific topics. Later they may be used for policy making, such as for defining financial actions in the 'investment agenda'. The monitor operates at a detailed level: the six digit (1234 XX) postal code area, rather than the borough level, for example. The monitor combines data on home ownership, income (purchased from Experian), 'WOZ-waardes', 'vroegtijdige schoolverlaters', age categories, households, ethnicity (Western/non-Western), 'schuldhulpverlening', WMO requests, etc. Furthermore, the goal is to also include police records . The monitor is a web application.

In time, the system may also be implemented in other cities. At this moment, collaboration with the cities of Purmerend, Dordecht, Eindhoven, Almelo is established. The aim is to develop guidance ('gebruiksrichtlijn') for the use of the system, including the notion of data minimization. Furthermore, no direct action can be undertaken based on outcomes from the system, it can only be used to support decision processes. It is not expected that completely new insights are gained, but that insights are obtained faster.

The pilot should be finished by July 2014.

2 Methodology

The case study focuses on the topic of liveability in the city of Rotterdam, and in particular three specific aspects: public safety, social cohesion, and physical environment (see 1.3). These aspects will not be addressed separately, but coherently. However, together they still represent a wide array of issues. As the purpose of this study is to gain in-depth insight into these developments, we map actors, their relation and the other findings drawing on a bottom-up research approach: Actor-Network Theory (ANT). Rather than strictly following this approach, we have used it as inspiration to draft our research methodology. This chapter presents the research approach, the specific topics that are addressed in this study, and the selection of the specific interviews and initiatives.

2.1 Theoretical background

In order to explore the data landscape around the theme of liveability in Rotterdam and its restructuring effects, this study draws upon Actor-Network Theory (ANT). ANT, which is a theory from the field of sociology, holds that all objects and ideas are socially embedded phenomena.[29] The theory enjoys a growing attention from researchers in the field of information systems, as it presents an alternative to technological determinism.[30] One of the main characteristics of ANT is the conceptualization of technology as one of the 'actors' in an actor-network analysis.[31] The notion behind this is that technologies are not just technological,[32] but they also possess human treats.[33] Technologies can thus embody social, political, psychological, economic, and professional commitments, skills, prejudices, possibilities, and constraints and objects may authorize, allow, afford, encourage, permit, influence, block, forbid and so on.[34]

ANT thus aims to investigate (technological) developments in a bottom-up manner and it treats technologies as well as humans as actors that have relations to other nodes in a network. Latour explains the use of ANT in practice as follows: "Follow the actors in their weaving through things they have added to social skills so as to render more durable the constantly shifting interactions".[35] This approach portrays society as a socio-technical web where technologies participate in heterogeneous networks that bring together actors of all types, human or non-human.[36] ANT particularly seems to have explanatory power in situations where innovations proliferate and where group boundaries are uncertain.[37] ANT thus appears to be useful for our study as we aim to investigate the data landscape in Rotterdam in a

[29] Latour, B. (2005). Reassembling the social, Oxford: University Press
[30] Latour, B. (1987). Science in action: how to follow scientists and engineers through society, Milton Keynes: Open University Press
[31] Walsham, G. & Sahay, S. (2006). Research on information systems in developing countries: current landscape and future prospects. Information Technology for Development, 12:1, pp. 7-24
[32] Bijker, W. & Law, J. (1992). Shaping technology/building society: studies in sociotechnical change, Cambridge, MA: MIT Press
[33] Latour, B. (2005)
[34] Latour, B. (2005)
[35] Latour, B. (2005)
[36] Law, J. (1992). Notes on the Theory of the Actor Network, Ordering, Strategy and Heterogeneity, Centre for Science Studies, Lancaster University.
[37] Latour, B. (2005)

bottom-up manner, to assess how data-driven innovation and the emerging data-ecosystem impacts liveability.

2.2 Research approach

Drawing on ANT, our research approach consists of four steps. Central to the approach of ANT is a network analysis. Therefore, we start out by looking at key actors and the way they deploy data and data analytics. Firstly we selected, in collaboration with the Ministry of the Interior, a shortlist of types of actors that make up the core of our investigation. To allow for a bottom-up analysis, the research approach starts by investigating who are the main actors regarding liveability, but also organisations that collect (potential) relevant data about the city. Secondly, based on interviews with the actors that are on the shortlist, we trace the actors they collaborate with (for instance providers of data or analytical tools) or who otherwise have an impact on their data-related activities (for instance regulators). Thirdly, in this way, using ANT allows us to map actors and their relation to each other; in short, to perform a network analysis.

2.2.1 Quick scan of initiatives and main actors

Together with the Ministry of the Interior, we selected a shortlist of the most important organisations and players from the public sector, the private sector and within the domain of the citizens. This selection was made based on desk research, media coverage or based on contacts in the network of either the Ministry or TNO. Examples include Rotterdam open data, Peerby and Iens as important online platforms that have data on the local level. This shortlist is provided in Annex 1. The methodology that we use are semi-structured interviews.

2.2.2 Full list of initiatives and organisations

During these first interviews, organisations they collaborate with were identified and these were subsequently interviewed (for the interview protocol, see Annex 2). Examples include data providers of Rotterdam open data. Hence, using a snowballing technique, the network of actors collaboratively forming the data landscape is identified. This resulted in a list of 33 organisations or experts that were either interviewed or about whom information was gathered (see Annex 3).

2.2.3 Network analysis

In order to perform the network analysis, the interviews were transcribed following the main themes of the semi-structured interviews: main objectives of data analytics, the process of value creation with data, interactions with other organisations, challenges and ambitions. Due to the scope of the project the interviews were coded by a single researcher, describing the relations in the interviews between actors: players, technologies and concepts. For the analyses of the network of actors the open source tool *Cytoscape* was used. An example of how this mapping of the relations between actors took place is shown in Annex 4.

The analyses resulted in several networks: an overall mapping of the ten most important actors in the interviews, and networks regarding the main themes of this study. These networks comprise somewhere between 10 to 30 actors to allow for a nuanced, but not too detailed description of the most important concepts. These networks are used to provide insights regarding the most important concepts, technologies and players, although players or only included if they were mentioned

by others during the interviews. The illustrations of the networks will be presented in the report in separate boxes next to the main findings and examples from the interviews. These accompanying descriptions of the networks will provide information of how the actors are related to each other, to the various types of players and the use of data and data analytics in the context of liveability.

2.2.4 Services, data and technologies

To complete the investigation, in addition to these network analyses that are based on the 'actors' in the interviews (players, technologies, concepts), several analysis were performed directly based on the interviews. These analyses score the types of data that are being used by the various players that were interviewed. Furthermore, the interaction between the various players will be used to sketch the data ecosystem of liveability in Rotterdam. These analyses support the network analyses described above.

2.3 Main topics during the interviews

The methodology used throughout the study is a semi-structured interview. This means that all interviewees are asked the same set of questions, but follow-up questions based on the interviewees' answers may differ. The mapping of the data-ecosystem focuses on a number of aspects, reflected in the interview questions (see Annex 2). The main themes during the interviews are: the objectives and position of the organisation in relation to liveability, the objectives and the process of value creation with data and data analytics (in particular whether they deployed profiling and personalization techniques), data, products and services that are produced and used, interactions with others in the network, and the ambitions, challenges and future developments. Some of these concepts are similar to concepts that Martijn de Waal uses in his work on 'mobile cities',[38] that also focuses on the relation between digital media technologies and cities (see Box 3). These themes were discussed during the interviews with organisations that deploy data and data analytics to support their core activities, organisations that provide data-services to other players, and organisations that do both.

BOX 3: Mobile Cities, example of relevant topics

'The Mobile City' is a dialogue between architects and city developers on the one hand and 'techies' on the other hand, focusing on mobile services. Often, smart city initiatives are based on the American – libertarian – notion of the extreme individualized society, such as propagated by Silicon Valley. At the same time this is not really in line with Europe's more communal notion of society. Currently, very interesting European initiatives take place in cooperation with architects aiming for re-development of buildings and areas. The link with digital media can be crucial, both in determining the purpose of a project as well as the development itself. De Waal's hypothesis is that digital media have a qualitative effect on such projects – and on the development of boroughs. He categorises online initiatives using four aspects: engagement, public, issue, and act. He sees a very strong tendency to increasingly profile on the level of postal codes – this has been going on since the fifties.

[38] Waal, M. de (2013). The City as Interface: How New Media Are Changing the City. Rotterdam: nai010.

3 The data landscape and use of data in relation to liveability

In this chapter we will look more closely at the data landscape of liveability in Rotterdam and liveability in cities in general. This analysis does not only include the actors who are directly involved in the process of providing and maintaining a certain level of liveability in cities and neighbourhoods, but also those that have data that are related to liveability. It will also highlight the most important concepts that frame the discussions on the deployment of data in the context of liveability and cities. Five aspects will be investigated in more detail: (1) the key actors regarding liveability, (2) the value creation mechanisms that lay underneath the deployment of data and data analytics and how this impacts value propositions, (3) the data applications and strategies that make up these value creation mechanisms, and (4) the main ambitions and (5) challenges that these actors have and encounter in relation to liveability. Subsequently, the next chapter four investigates the emergence of an actual data ecosystem around liveability in the city.

3.1 Key players regarding liveability

Many actors are involved in ensuring liveability. As discussed in the previous chapter, we used a bottom-up approach, combined with a network analysis to investigate the emergence of a data ecosystem. The organisations that were interviewed as part of the investigation are shown in Table 1 (note that organisations can be mentioned multiple times).

Table 1: Overview of the organisations that participate in this study

Traditional role	**Service with direct impact on liveability**	**Service with indirect impact on liveability**	**Supporting services**
(Local) government	- Stadsbeheer - Police - Buurt Bestuurt		- Safety Index - Leefbarometer - Open Data Portal
Semi-public	- SOR - Achmea - Plan B		- Veldacademie - Hogeschool Rotterdam - TNO
Commercial organisation or service provider	- 2CoolMonkeys (*Bomenspotter* app) - Peerby	- TomTom - OMA/AMO - Funda - Iens - Marktplaats - Eneco	- Google - TomTom - IBM - Esri - Thingful - Sense-OS - EDM - StartupBootcamp - GfK - 2CoolMonkeys - KPN - Sanoma

The organisations that were interviewed as part of the investigation are categorised into public, semi-public and private organisations according to their primary task. Furthermore, they are categorised according to their relation to the services that are provided in the domain of liveability:
whether these organisations are directly involved with liveability, e.g. with their products and services, indirectly (e.g. because they collect data about the city and deploy these for services that are indirectly linked to liveability) and supporting services that currently provide data services (e.g. analytics) to the first two categories.

Key actors in the field of liveability in the city are (local) governments (e.g. the municipality of Rotterdam and the Ministry of the Interior and Kingdom Relations) and governmental organisations such as the (national) police and Stadsbeheer (the municipal organisation responsible for the maintenance of the public sphere, such as maintenance of physical objects and waste disposal. In addition, the municipality has set up public initiatives such as the *Buurt Bestuurt* ('Neighbourhood Governs') initiative that aims to improve social cohesion and safety in neighbourhoods by facilitating direct interaction between citizens and professionals such as the police. Furthermore, semi-public organisations, such as health insurance company Achmea and housing corporation SOR, which focuses on housing for the elderly, compile data on cities as part of their respective public tasks. Some commercial service providers have a direct link with liveability too. For instance, apps such as *Beter Buiten* ('Better Outside') which allows citizens to report broken public objects, the more leisure-focused *Bomenspotter* (Tree spotter) app, and *Peerby*, a website that facilitates the sharing of products between neighbours, which could foster social cohesion.

Other organisations have an indirect relation with liveability, either because of their local presence (e.g. a retailer) or because they collect data about citizens and their interactions with specific locations, often in a specific domain, such as *Marktplaats* (the largest Dutch online marketplace, owned by eBay), *Iens* (the largest Dutch online directory for restaurants and bars), *Funda* (the largest Dutch online directory for houses), KPN (the largest Dutch telecom operator), Eneco (one of the largest Dutch energy companies), and global players such as Google, Twitter and Facebook and TomTom (provider of navigation products). Also, citizens are not mere consumers of governmental and private services. They can also participate in the various steps from agenda setting and policy design to execution and evaluation. What is notable, however, is that no (semi-)public organisations that took part in this study have an indirect role in relation to liveability. These organisations either take on a direct or a supporting role.

Figure 1 depicts the content of Table 1 in a graphical manner. By showing the different roles of the organisations that took part in this study, a first outline of a value chain emerges. The role of citizens as not mere consumers, but also as participants, is depicted by the two arrows. Not all organisations could be fitted in the picture. Therefore, we show examples of organisations that take up a specific role in the value chain.

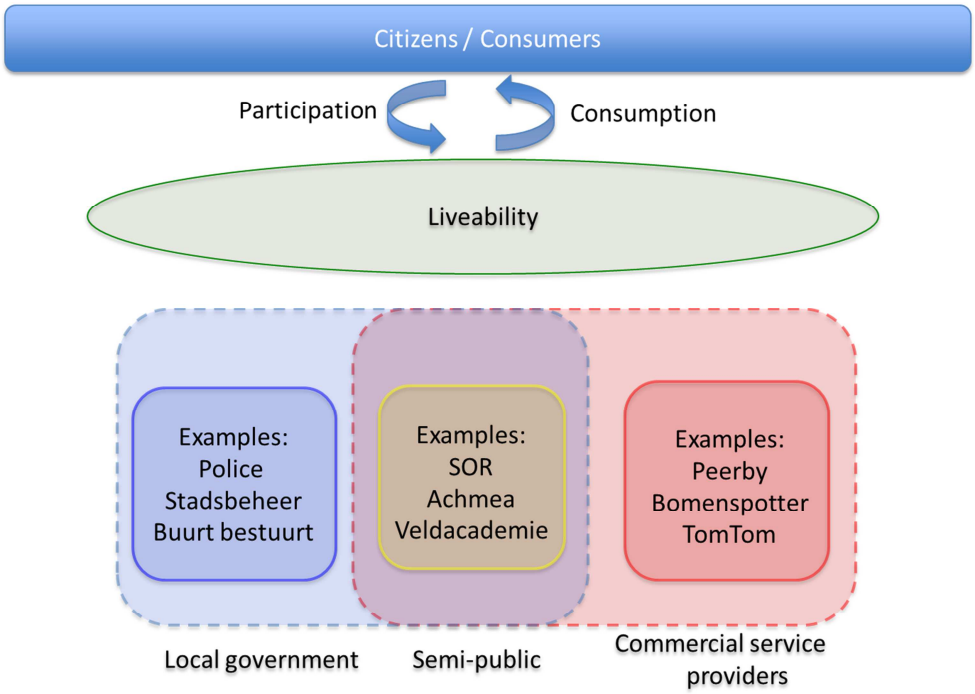


Figure 1: Value chain of liveability

The collection, analysis and presentation of data has always played a part in the value chain of liveability. However, due to the 'datafication' of our society and the advent of smart cities this is expected to increase. In addition to organisations that have a direct link to liveability, there is a myriad of organisations that collect – often via the activities of their users – vast amounts of data about cities and neighbourhoods, which could potentially be valuable in the context of liveability as well. Consequently, these service providers could be relevant players in the field of liveability in cities. An interesting example was provided by Eneco, one of the largest energy providers in the Netherlands, serving over two million households. An Eneco spokesperson stated, after the news broke out that an elderly woman passed away and was only found in her home in Rotterdam after ten years, that these kinds of incidents could be detected based on patterns in data on energy consumption.[39] This example illustrates that data collected by providers of utilities, infrastructures and products and services are potentially part of the same ecosystem of data in the city, even though the delivery or maintenance of services directly impacting liveability is not their main concern.

Therefore, the next section investigates which value creation mechanisms exist and emerge around data collection, analysis and use, and how this relates to the domain of liveability.

[39] Kamerman, S. (2013) De mensen hebben geen contact meer met elkaar. In: *NRC Reader*. Available at: http://www.nrcreader.nl/artikel/3316/de-mensen-hebben-geen-contact-meer-met-elkaar

3.2 Value creation mechanisms of data and data analytics

In general, it is possible to distinguish four phases in the development of a data strategy within organisations that (want to) deploy (big) data analytics: efficiency, effectiveness, new propositions and transformation.[40] Whereas the first two phases focus on the deployment of data analytics to improve efficiency and effectiveness to enhance current value propositions of an organisation, the third phase signifies the development of new value propositions that are enabled by data analytics, for instance by providing personalized products or data as a service. The fourth phase is an extension of the emergence of new value propositions, as the redefinition of value propositions could enable an organisation to take a strategic position as markets are being restructured due to their datafication. For the analysis of value creation mechanisms this study, however, we combine these four phases in two categories: the enhancement of existing value propositions and the development of new value propositions (see Table 2 and Annex 5).

An analysis of the value creation mechanisms and the resulting value propositions of data and data analytics (see Table 2 and Annex 5) reveals that currently data analytics is most often deployed by the organisations that were interviewed to facilitate or *enhance* internal (production) processes. This applies to both commercial (online) companies and governmental services. The most common *new* value propositions that were found in this study enable the above-mentioned enhancement of the primary process and business model. They are usually business-to-business oriented and delivered by information or data analytics providers and platforms. New propositions for consumers or citizens are far less common. They comprise personalised, adaptive (information) services, smart devices, wearable technology and smart cars, although they were often spoken of in terms of future developments or ambitions. We will elaborate this in more detail below.

Table 2: Data-driven value propositions

Organisation	**Enhancing existing value propositions**	**New value propositions**
Google		
TomTom		
Iens		
Achmea		
Gemeente Rotterdam		
Eneco		
KPN		
Police		
Stadsbeheer public sphere		
Stadsbeheer safety		
Marktplaats		

[40] Veenstra, van A.F., Esmeijer, J., Bakker, T., et al. (2013) Big Data in small steps: assessing the value of data. Available at: http://publications.tno.nl/publication/1106495/46DP56/veenstra-2013-big.pdf

Organisation	Enhancing existing value propositions	New value propositions
Funda		
Peerby		
Sanoma		
SOR		
Municipality of Almere		
OMA/AMO		
Leefbarometer		
Veiligheidsindex		
TNO-Ducha		
BuurtBestuurt		
Startupbootcamp		
Plan B		
Umbrellium		
Sense-OS		
Esri		
2CoolMonkeys		
IBM		
EDM		
GfK		

Policy makers and municipal organisations such as Stadsbeheer, for instance, use data to create management information, just like companies such as Eneco, Achmea and retailers. In the city of Rotterdam, policy makers rely on data to design, monitor and evaluate their policy in terms of public safety, maintenance of the public sphere, housing services and development and the impact of certain interventions. In addition to this long-term focus, data is also being deployed for the daily operations of various municipal organisations, such as Stadsbeheer, which uses data to efficiently perform the maintenance of the public sphere (physical objects such as traffic signs, trees, street lighting, etc.), waste disposal and safety on the street. Although these operational objectives are the most important focus, occasionally Stadsbeheer deploys data to address very specific questions that are relevant for these operations. For instance, it has combined data regarding mutations in the field with demographic data from the Gemeentelijke Basis Administratie (municipal citizens' administration) to assess why certain objects on a playground will be vandalised because they may not be suitable for the average age of the population in the neighbourhood.

The police uses data to support their teams dealing with incidents and emergencies, for briefings of police officers that patrol the streets (which also entails a form of predictive policing), for investigation and occasionally for crowd management. In case of an emergency call so-called Real-time Intelligence Centres search for information in police databases and online sources to provide the emergency team with relevant information. In the field of social cohesion, the *Buurt Bestuurt* initiative that supports some neighbourhoods by setting up committees

consisting of citizens and supporting professionals (such as police officers, policy makers, community workers) focuses on the most important safety-related issues in the neighbourhood. In order to foster interaction between professionals and citizens there are physical meetings and citizens can use the *Buurt Bestuurt* app to vote for the most important issues that should be addressed.[41]

Marketing and customer interaction are fields where data analytics proliferates, particularly within commercial companies. For online services the use of data is more often a crucial part of the service it provides. In case of online services such as provided by Google, Funda, Iens, Peerby and Marktplaats data is at the core of the organisation ensuring that the service functions, matching supply and demand of advertisements (Google), houses (Funda), restaurants (Iens) or products to share or sell (Peerby and Marktplaats). Other applications of data analytics are monitoring the website performance, A/B-testing of adjustments to the website and communication strategies and advertising. TomTom uses a signalling system to automatically detect anomalies in traffic patterns to see whether their maps need to be updated (which, subsequently, is done manually to ensure quality and control). Marktplaats uses a signalling system to detect anomalies in user behaviour that may indicate a problem with the website. Iens uses tools that support its editors to automatically spot user contributions that require attention (e.g., by filtering on specific keywords). Iens also uses data analytics to integrate and weigh user reviews of restaurants: reviews from experienced contributors have a bigger impact on the overall rating of a restaurant than the review from a novice contributor – in a way profiling users based on their reviewing-track record. These innovations are related to the current business model and existing products and services.

As mentioned above, the deployment of data analytics in existing organisations does often not result in the development of new products or services. Even though many service providers collect data that could be interesting for others, commercial exploitation of data or intelligence as a product is not common practice yet. For example, while Rotterdam Open Data provides a lot of data from the municipality, it currently provides no additional analytics to enhance the data. Some exceptions exist, such as TomTom and Achmea. TomTom uses data to add real-time traffic information to its navigation service. Furthermore, it sells intelligence derived from its traffic data to third parties. For instance regarding the use of infrastructure to municipalities or analyses that provide insights to, for instance, retailers or transportation hubs like airports that would like to know – on a aggregated level, where travellers to the airport travel from. Achmea uses data to develop benchmarks for health providers and professionals enabling them to improve their service level and the organisation provides researchers with data for academic purposes. Other organisations occasionally use data to generate insights for third parties as part of their PR strategy. Marktplaats, for example, occasionally provides overviews of the most active barter-municipalities in the Netherlands, but these are primarily tailored to (local) media rather than policy makers. Providing software as a service to third parties is also rare. Google offers several services (even cloud-infrastructure), and so does TomTom (e.g. fleet management). Iens has started to

[41] In Almere and Eindhoven, pilot projects are undertaken that are much more data intensive. In Almere, the idea of an early warning system is currently being piloted, with a different focus: wellbeing and social security. The idea is to support social- and community workers by detecting patterns in data that indicate potential social issues that could be addressed in an early stage.

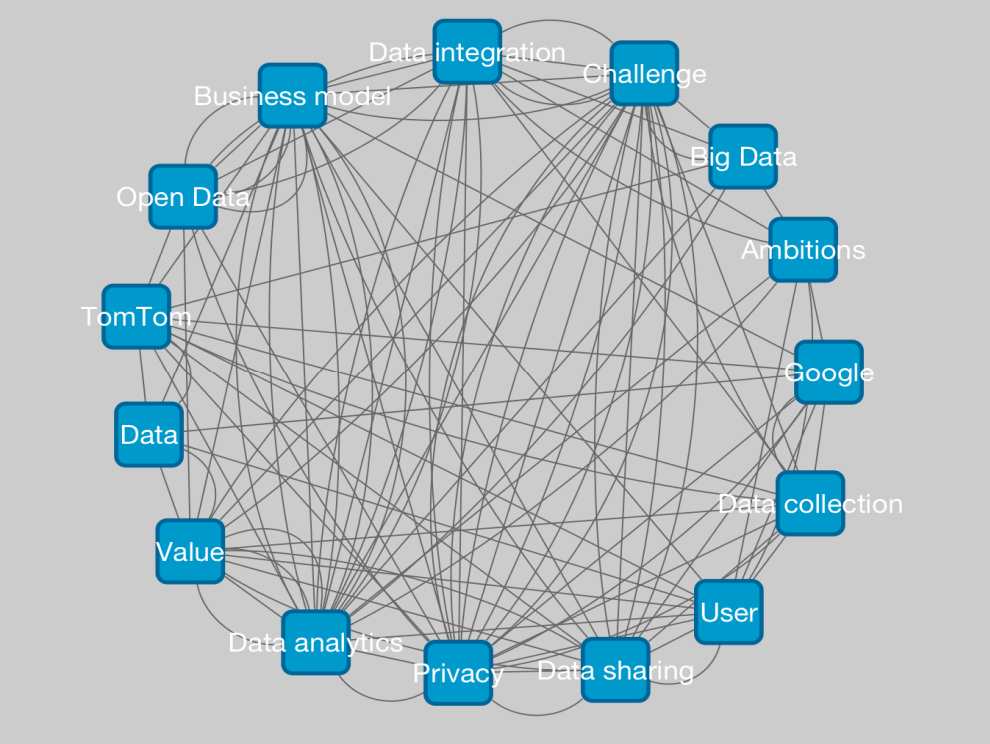
integrate a reservation system for restaurants into their website and offers a website design service.

Some semi-public organisations and commercial organisations also deploy data for objectives that are related to liveability because this is part of their core business, or because they have developed new value propositions that (indirectly) concern aspects of liveability. Housing corporations, for instance, use data to manage their offering and for matching supply and demand. Insurance company Achmea provides health providers and professionals with benchmarks (based on data from areas with a similar demographic situation) that enable local health service providers to improve their service, and TomTom sells information about the use of infrastructure to municipalities. Some retail chains and supermarkets use data to determine whether and where they would like to expand although their decisions are driven by revenues rather than the aim to improve liveability in neighbourhoods. Some service providers use data to create new applications that can be used by citizens to enhance the liveability, such as the *Bomenspotter* app from 2CoolMonkeys.

The way data analytics is deployed in an organisation, this study found, is strongly determined by existing business models and practices in these organisations. The graph in Box 4 presents the most frequently cited concepts in our interviews on data use and liveability, and their correlations (each line between two concepts represents an explicit relation between them that was mentioned during the interviews). Further analysis reveals that interviewees are first and foremost concerned with valorising the use of data analytics; it is most strongly connected to the concept 'business model' (these concepts were linked to each other six times during the interviews), 'value' (four times) and 'data integration' (four times), although the latter is more concerned with the 'what' and 'how' of data analytics.[42] Data analytics is expected to provide insights that support the daily operations, enforce the current business model (or public task) and, subsequently, generate value to the organisation. This mechanism is the driving force in the strategy of data analytics and explains, to some extent, how organisations deploy data and data analytics.[43] This focus on current business models and practices explains the current focus on the enhancement of current value propositions, rather than on developing totally new value ones (see Box 5 for more information).

[42] Box 4 also illustrates how privacy is the most important challenge (7), followed by data sharing (4), data integration (3) and actually generating value from data analytics (3).
[43] In addition to the business model as the driving force of data strategies, part of the data analytics is – to some extend – driven by legal obligations in terms of accountability (e.g. insurance company Achmea and energy company Eneco are required to provide information to public authorities for compliance monitoring).

BOX 4: Most cited topics during interviews



BOX 5: The influence of business models on the application of data analytics

The influence of the business model on the use of data analytics can be very nuanced. Funda, for instance, makes money through selling advertisements, but its main customer (and shareholder) is the NVM, the Dutch Association of Real Estate Brokers, that uses Funda as a tool to direct potential buyers to the brokers who facilitate the transaction. It is currently working on the design of a recommendation system, which serves both the user and the broker, but it does not display all available information on their website (for instance crime rates in a neighbourhood) because that is not in the interest of the broker. Consequently, Funda always needs to find a balance between serving the user of the site and the NVM. Furthermore, Funda does not pursue revenue maximisation and consequently it is not looking for ways to further exploit the data.

Another example is Marktplaats (which is owned by eBay) where people can sell and buy products from small vendors. The development of a recommendation system, which includes some form of predictive analytics, could be interesting to Marktplaats, but it is not a priority because the organisation does not make money from the transactions (these take place offline), but through selling advertisements. Because they do not possess transaction data, (predictive) recommendations are more difficult for Marktplaats than for companies that do facilitate a transaction (such as Booking.com or KLM). Another example where predictive analytics fits the core activity of an organisation is TomTom's satnav system, which includes both historical and real-time traffic information and to direct the users of the system and adjust their navigation services if necessary.

3.3 Data applications and data strategies

In the previous section we investigated the value creation mechanisms of data and data analytics and the resulting value propositions of organisations deploying data. This section explores the ways in which organisations deploy data and which data strategies they have to create value with data. Looking at the application of data by the organisations that were part of this study, on a generic level two major applications of data can be distinguished: (1) an increase of automation based on data or automated decision-making ('data-driven action') and (2) an increase of contextualization or personalization. While these trends can be separated analytically, they can coincide. In case this happens, they lead to 'customized action'. The organisations and the data-driven initiatives that were investigated in this study, are categorised in Figure 2. Some organisations deploy multiple data strategies.

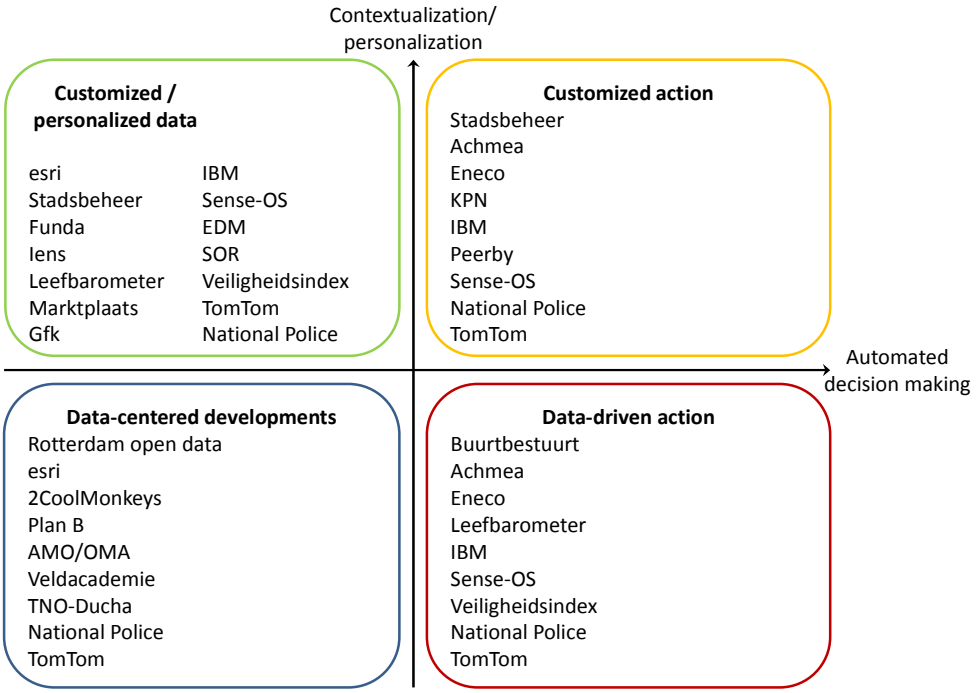


Figure 2: Data strategies of the organisations that participate in this study

As shown in Figure 2, organisations take different approaches to deploying data and to developing data strategies, but some patterns emerge. Contextualisation and personalisation are deployed by more organisations than automated decision making. Geo-information company esri, for example, is able to map any information onto a very detailed and precise map. Marketing organisations Gfk and EDM assemble customer profiles to tailor retailers strategies to postal code areas, or even to specific customer profiles. Geographic information is currently mentioned as the most interesting way of contextualising, followed by personalisation in case of organisations that have a legitimate reason for processing personal data, or that are able to assemble information into generic customer profiles. Other ways of contextualising information, such as making data 'social' by adding social media data or making data 'real-time' by basing information on sensor data are also

mentioned often, but usually as a next step in the development of data-driven innovation.

Based on the overview of the activities that the organisations that have taken part in this study shown in Figure 2 another observation can be done. Organisations that focus on developing a (standardised) infrastructure, such as Stadsbeheer, Sense-OS, IBM, the National Police and TomTom are more easily able to develop services that represent both trends than organisations that only focus on the development of specific services. Similarly, it can be expected that organisations that develop a platform that is used by more actors or even by the sector as a whole may become the new key players in the future of the data landscape. Currently, more organisations attempt to become such a player within a wider area than their own business, such as KPN, Sense-OS and TomTom. Others, such as Stadsbeheer and the National Police, mainly aim to become platforms for their own policies and tasks, while IBM focuses explicitly on developing the technology, hardware and services for their clients to develop such services and platforms.

3.4 Ambitions

In the governmental domain and among commercial service providers there is an ambition to improve and expand analytical activities beyond improving current operations (see Box 6 for an overview of the most cited concepts related to the ambitions and how they relate and Table 3 for a categorisation of the ambitions mentioned by the organisations taking part in this study). Although analytics for the contextualisation of advertising – predictive and personalised based on profiling technologies – are not yet as commonly deployed by the organisations that have been interviewed for this study as the deployment of data analytics for the current operational activities, there is consensus that the use of data has great potential.

BOX 6: Ambitions

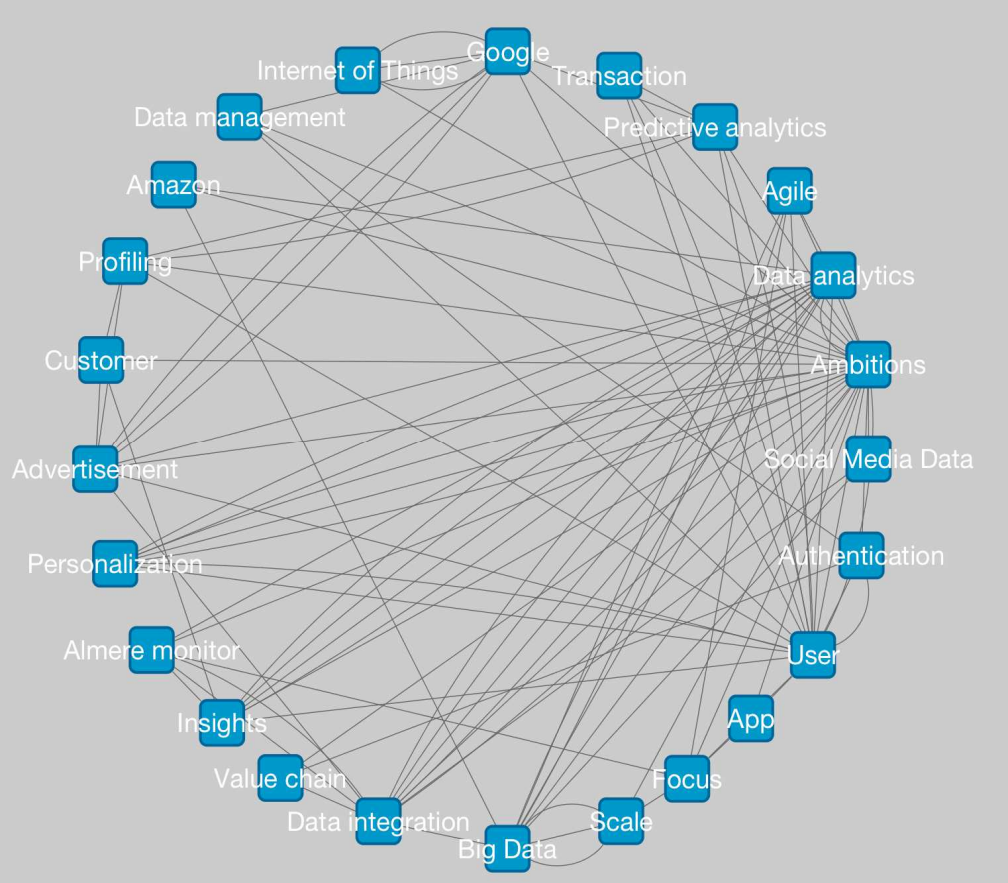


Table 3: Ambitions mentioned by the organisations that take part in this study

Category of ambitions	**Types of ambitions**	**Number of organisations**
User participation	Interaction with consumers or citizens, the use of social media	15
Contextualisation	Profiling, personalisation, predictions	15
New products and services	Innovation, new business models, data as a product, intelligence as a service	13
Internet of Things	Including sensors, smart city systems	9
Improving operations	Efficiency, decision making	9
Visualisation and analysis	Presentation, image analysis	9
Communication	Marketing, communication	8
Transparency	Transparency, democratisation	6

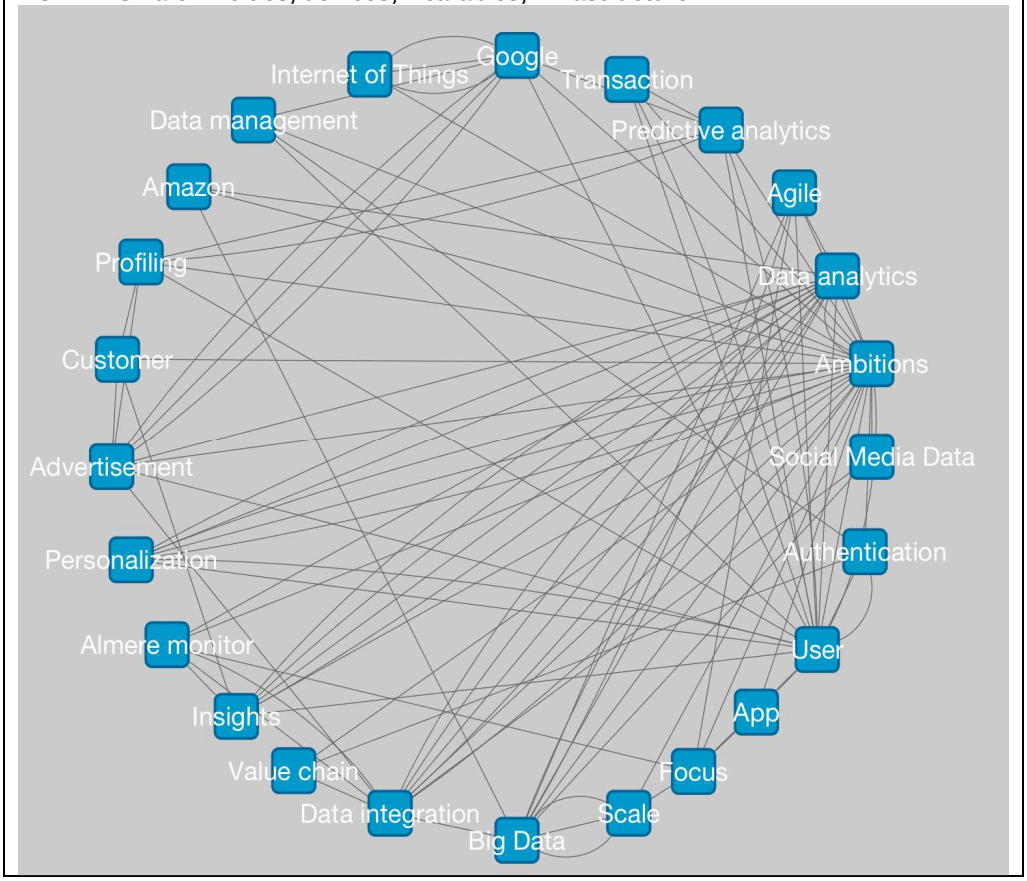
Table 3 lists a number of ambitions that were often mentioned by the organisations that took part in this study. The three ambitions that were most often mentioned were (1) user participation and making services ‘social’ by the use of social media and by facilitating interaction with consumers or citizens, (2) contextualization by

deploying profiling techniques, predictive analytics and, consequently, personalisation, and (3) the development of new services. An examples of the first is the ambition of the Rotterdam Open Data portal to allow feedback on the datasets in their portal, while examples of the second include Funda and Marktplaats, which both aim to develop predictive capabilities based on their current services and the data that is used in these services. Notable is the third ambition. While many organisations claim to mainly use data to improve their current business model, they clearly have the ambition to develop new products and services. These ambitions are also reflected by the result of the keyword analysis in Box 6.

Another ambition of many organisations is connecting to the IoT and becoming connected to systems of the smart city. Nine organisations mentioned this. With the development of wearable technology, such as smart phones and smart watches, in combination with the smart home and smart grids, a new range of services seems to emerge. The integration of these various 'smart' spheres and their data is expected to provide the infrastructure for smart city developments. Especially wearables and cars were mentioned in examples of platforms for distribution of contextualised information. However, although more and more sensors, devices and machines are getting connected it is still too soon to see an actual global (or even local) IoT. According to one of the interviewees the current state of the IoT is similar to the early days of the internet before the introduction of commercial web browsers when people were not aware of what was going one beyond the local community they were active in.

Although the IoT and sensor data, thus far, is only sparsely integrated and deployed in the domain of liveability, the data that are generated about the city and its citizens seem to hold great potential (see Box 7 for the most cited concepts related to this new field). Especially Google and to a lesser extent TomTom (which has developed a smart watch) are active in this domain. Google recently acquired several robotics companies and Nest, a smart meter and home security system, which signifies – together with its *Android for the home* operating system, a first step in the connected home. Microsoft, Samsung are also active in the domain of the smart home, as well as home automation companies. Energy companies are another type of player active in the home with their smart meters and the advent of smart grids. Data sharing and aggregation (see Box 7) play an important role to realise these ambitions. This requires integration of formerly siloed databases within organisations, for instance transactional data and browsing history, but also among organisations, such as in the case of acquiring social media data.

BOX 7: "Smart"… cities, devices, wearables, infrastructure



3.5 Challenges

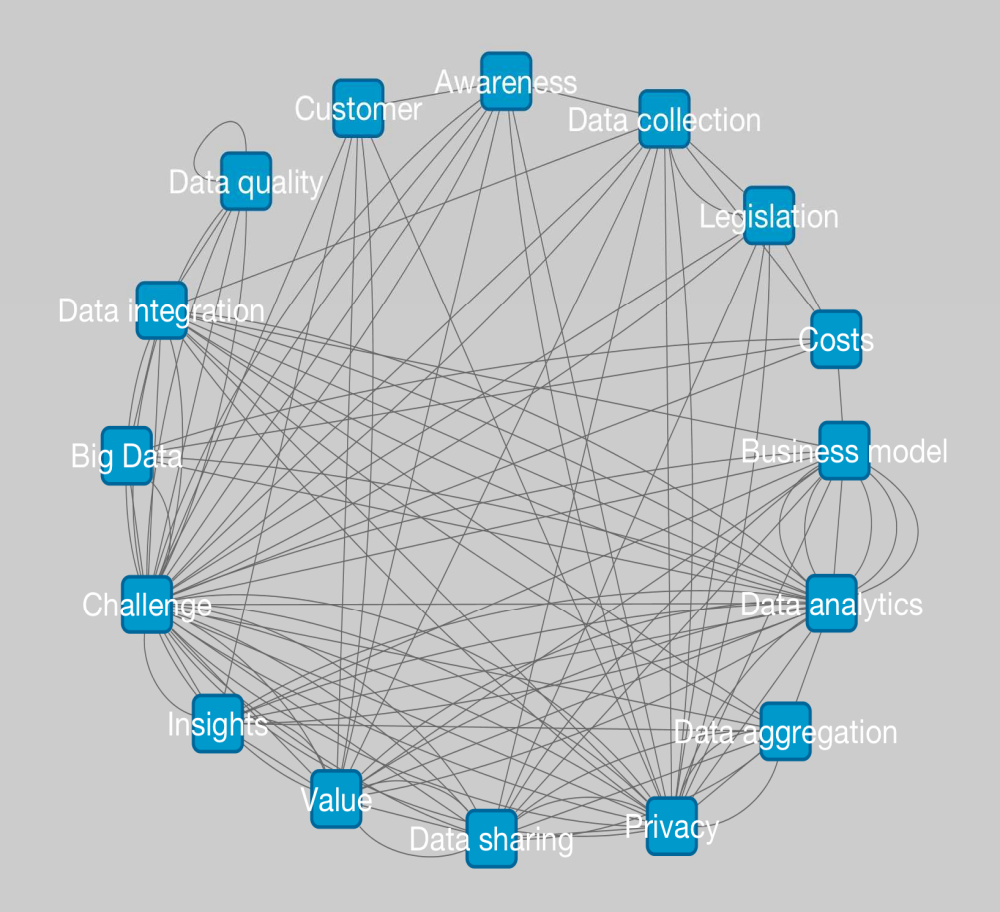
While deploying data and data analytics is expected to yield new value, it is not without its challenges. The most important challenges (as counted in Table 4 and illustrated by the number of links between the concepts in Box 8) are related to privacy (which is related both to actual challenges regarding compliance with legislation and to perceptions of privacy), organisational readiness and awareness of the potential of data for new services, the actual process of successfully providing new and relevant insights, and data quality.

Table 4: Challenges mentioned by the organisations that take part in this study

Category of challenges	Types of challenges	Number of organisations
Privacy	Privacy, data protection, use of personal data	18
Organisational readiness	Awareness, organisational culture, and organisational skills	18
Data quality	Data quality	13
Value mechanisms	Link with value propositions, (new) business models	12
Technology	Data integration, standardization, legacy systems	12
User involvement	User participation, interaction with citizens	6

A lot of the datasets on liveability in the city contain personal data which means privacy is a first major challenge, especially when multiple datasets are integrated and used for new purposes. The Early Warning System in Almere, for instance, tries to collaborate with the Dutch Data Protection Authority (CBP) as it treads a fine line by combining several datasets to create profiles on a six digit postal code area level, in which it also looks at demographic characteristics. It illustrates that at its core one of the qualities of big data is that it can provide information with a high granularity (creating profiles that hoover somewhere between the level of individuals, demographics, social groups and street level).

BOX 8: Main challenges cited in the interviews



Another aspect of data-driven innovation and big data is that combinations of datasets may lead to new insights. However, there are legislative restrictions – designed to protect the privacy of citizens – regarding personal data or the combination of datasets that could lead to information that can be used identify citizens. One of the interviewees described how this makes it difficult to ask for consent from users to use their data to develop a new service based on the use of data, because nor the organisations aiming to process personal data, nor the users cannot properly determine the information that can be obtained by combining data sets. And privacy is not only a matter of legislation. Several interviewees emphasised that organisations not only have to deal with the law, but also with

customers' perceptions. One of the interviewees referred to the public backlash when Equens, an organisation handling payment data for banks, presented its plans to sell aggregated information based on the transaction data of consumers' payments.[44]

A second challenge concerns the awareness of organisations of the (potential) value of their data and how this could be deployed to develop new insights, products or services. As described above, data analytics is often deployed at the core of the business model. New applications of data analytics require time, human resources and money that are not directed towards the primary process of an organisation. Therefore, awareness demands a certain belief because data analytics requires an investment in staff, tools, infrastructure and possibly new data sources. Subsequently, awareness needs to be translated into commitment, especially by higher management.

A third challenge is the deployment of data analytics, which requires special skills. Especially when organisations engage with analytics that extend their traditional business intelligence activities by asking new kinds of questions (for instance regarding profiling and predictions) and combining new kinds of datasets, new skills are required. The concept of 'big data' is explicitly mentioned in the interviews in this context. Furthermore, several interviewees indicate, as the complexity of analytics increases, the importance of a certain level of democratisation of the data analytics, actively engaging domain experts by making the analytics mode accessible, for instance with visualisations and 'playgrounds' where they can use data and business intelligence tools to define new, more complex queries that can be executed by the actual data experts.

A fourth issue, the quality of the data, is directly linked to data integration and analytics. While data quality always represents a multi-faceted issue (e.g. regarding the integrity of data, frequency of updates), for example when multiple organisations collaborate, especially the integration of citizen and user participation in formal processes and databases can be difficult. These difficulties are both technical and organisational. Stadsbeheer, for example, uses data collected by citizens via the *Beter Buiten* app to detect mutations in the field, for instance due to vandalism. However, this is not directly processed in the Stadsbeheer administrative back office, because they need to maintain quality. The police use data from Twitter and the more formal Police app, but this data is not part of the formal police reports that, for instance, will be used in court. The police works with a rigid chain of command and wants to keep control and maintain their existing protocols. This is also linked with authentication and a better understanding of the value of information from other sources, such as citizens/users. As described earlier, for Iens – where user participation is core – reviewers need to log in before they can post something and Iens takes the experience of the reviewer into account to process the review in the overall rating of a restaurant to maintain the overall quality of its service.

[44] http://www.volkskrant.nl/vk/nl/2680/Economie/article/detail/3446572/2013/05/24/Equens-ziet-voorlopig-af-van-verkoop-pingegevens.dhtml.

4 Towards a data ecosystem

The previous chapter introduced the outlines of the data landscape regarding liveability in Rotterdam, including the organisations that are active in this field, the value creation mechanisms of the data analytics behind their existing and new value propositions, their data applications and data strategies, and their ambitions and challenges regarding data-driven innovation. The third step of this study is the description of the emerging data ecosystem in the domain of liveability in the city of Rotterdam in order to investigate its impact on and implications for this domain.

4.1 The data ecosystem of liveability

Data in and of itself does not represent much value. In order to transform raw data info actionable insights, a value creation process is required. There are numerous depictions of the value creation process of data[45] [46] [47] [48] that each build on its most basic form which can be represented in three main steps: (1) data generation, (2) data analysis and (3) output. Data-driven innovation is the result of the concrete fulfilment of this value creation process of data in order to reach a specific goal; to tackle a problem or grasp an opportunity to which data analytics could provide (a part of) the solution. For each specific goal, an organisation (or a consortia of organisations) needs to organise a value creation process. It is likely that for many of the steps in this process organisations will have to involve third parties, because they lack experience, technological resources or talent to deal with the multidisciplinary aspects of big data on their own. The resulting network of actors is in most cases specifically tailored towards the goal that is being pursued. The aggregate collection of big data players, services and products, technologies and regulations that shape the data-driven constellations create a transcending (big) data ecosystem.[49] As described in chapter one, the perspective of the business ecosystem allows us to not only describe practices, ambitions and challenges from individual organisations, but also their interactions in the broader, multi-faceted context of data in the city.

Before we are able to describe the emerging data ecosystem of data in the city in this chapter, based on the analyses in the previous chapter, we discuss the most important findings regarding the interactions between actors that play a role in the fulfilment of the value creation process in the domain of liveability in Rotterdam.

---

[45] Manyika, J., et al (2011). Big Data: the next frontier for innovation, competition, and productivity. McKinsey Global Institute. Available at: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation

[46] OECD (2013) Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data", *OECD Digital Economy Papers*, No. 222, OECD publishing. Available at: http://dx.doi.org/10.1787/5k47zw3fcp43-en

[47] Tech America Foundation (2012). Demystifying Big Data: A Practical Guide To Transforming The Business Of Government. Available at: http://www.techamericafoundation.org/bigdata

[48] Kaisler. S., et al. (2013) Big Data: Issues and Challenges moving forward. In: *IEEE Computer Society*, pp. 995-1004.

[49] Esmeijer, J., Bakker, T. and Munck, de S. (2013) Thriving and surviving in a data-driven society. TNO report R11427.

4.2 New insights require data integration and data sharing

The figure in Box 4 illustrates the importance of data integration in the domain of data in the city. It signals a prevailing awareness that the combination of datasets could yield new valuable insights. As organisations aspire to do more with data analytics, data integration is high on the list of ambitions (see Box 6), while it is also one of the most important challenges related to the actual analytics (see Box 8). For some organisations data integration means combining datasets that they already collect, but store in separate siloes, such as online user behaviour and transactional data, or online data and offline behaviour in the case of retail. However, in some cases data integration also refers to the combination of one's own data with data from third parties, which can be open or proprietary. This implicitly refers to all kinds of (potential) partnerships. For example, as mentioned above, Stadsbeheer combines data collected by its fieldworkers with the administrative database to create an overview of the state of the physical objects in the public sphere in Rotterdam. But it also includes data from citizens that are reported by the privately owned *Beter Buiten* app. However, Stadsbeheer does not collect data that is produced by citizens via social media. In that sense, the 'physical world' (what happens) and the 'social world' (how people feel about what's happening) are currently not connected by Stadsbeheer.

Police officers actively engage with citizens on a regular basis, both by physical encounters in the neighbourhood or digitally via a special Police app or social media, especially Twitter and Facebook. The data collected via the Stadsbeheer app is also made available to include in their daily operations. In case of an emergency call there are so-called Real-time Intelligence Centers who collect all relevant information from police databases and online sources to provide the emergency team with all the relevant information. Other data sources that the police use are the municipal citizens' registration, the European CAR and driving licence Information System EUCARIS and its Dutch counterpart at the RDW, the intergovernmental Schengen Information System (SIS), which contains data about persons or property of interest, the Kadaster (national houses and property registry) and its own administration database, and it is also incorporating the New Trade Register containing all commercial organisations. In some cases (e.g. detection) the police can also gather proprietary data, for instance CCTV-images from shops. In that sense, the police is one of the exceptions as it explicitly integrates various external datasets to improve their own information position.

The work of policy makers and more long-term oriented governmental operations rely on several monitoring tools, which are not based on (near) real-time insights, but rather on developments over a longer period of time to assess the impact of policy and interventions, and design new ones. The municipality of Rotterdam has several instruments that are being used to this end, often reports (for instance from the research organisation *Veldacademie*) or biannual monitoring tools, such as the *Safety Index*, the so-called *Social Index* and the *Physical Index* to create a single monitor for an integrated Neighbourhood profile, and the *Housing Service Monitor* (Woonservicemonitor). On a national level the Dutch Ministry of the Interior and Kingdom Relations has developed the *Leefbarometer*, which is updated every two years and scores the level of liveability on a neighbourhood-level and is aimed to support municipalities to evaluate and design their policy. The *Leefbarometer* is based on demographical data from the Dutch Central Bureau of Statistics (CBS),

data on houses and property from the Kadaster, data on rehousing from Cendris and consumer and lifestyle data from BisNode. Other long-term developments, which are not so much focused on real-time insights, are housing and infrastructural projects. For spatial planning in Rotterdam data from organisations like Veldacademie or urban thinktank OMA/AMO are used. Some cities acquire data from TomTom to monitor and improve their roads.

One of the most cited value creation mechanisms of data analytics regarding liveability (see Annex 5) is that it enables more integrated interventions. This integrated approach was already more common in policy design, but as the examples above illustrate, this notion is becoming more important in daily operations as well. This could signal an increasing integration of multiple (real-time) datasets from governmental services, commercial services and citizens, which demands new partnerships, that are currently not yet in place.

4.3 Services use geographical data most, next are social and sensor data

The use of certain types of data is depicted in Figure 4. Based on the interviews, we mapped the types of data that organisations possess to a few categories: 'traditional' transactional data, geo-graphical data, sensor data, social media data, media content and personal data. Naturally, organisations could possess use of multiple types of data. This Figure shows that, currently, most combinations deployed by both governmental and commercial services focus on transactional data and geographical data. Another important type of data is social media data, for instance data from Twitter, Facebook or Instagram, especially for marketing purposes, which is often closely linked to the business model. However, the integration of social media is still in its infancy and not yet broadly adopted by the organisations that were interviewed. Another data source often mentioned among the ambitions of organisations is the use of sensor data, which is also reflected in Figure 4.

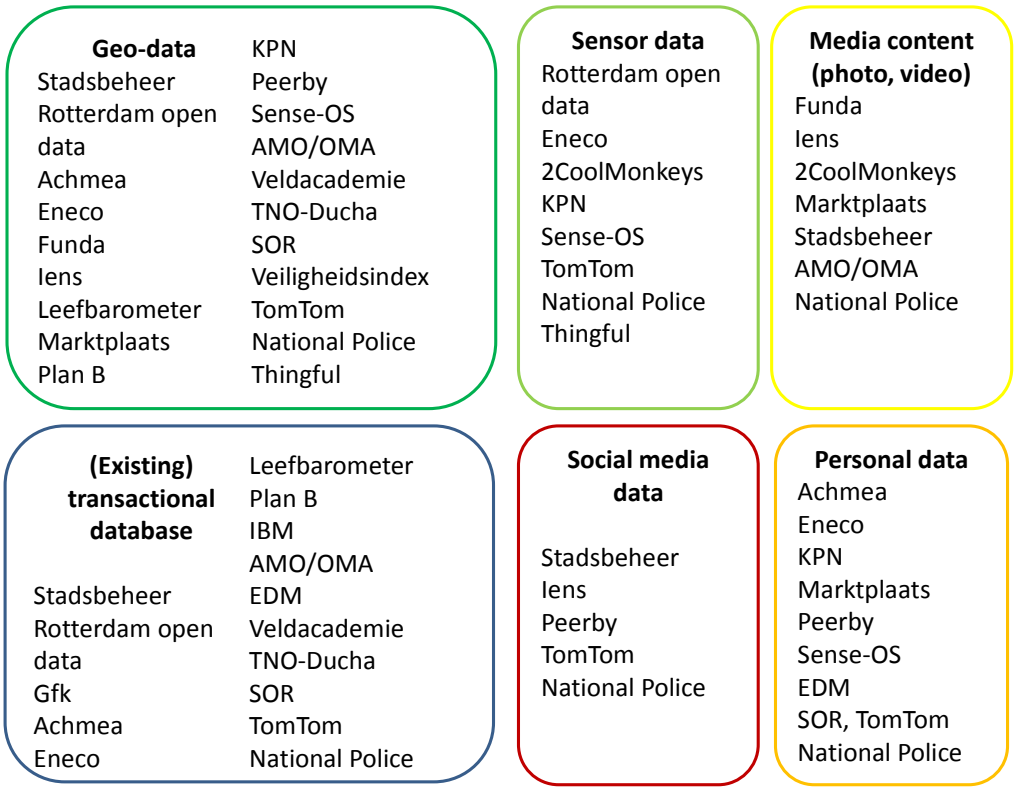


Figure 4: Types of data in the city

4.4 Data platforms facilitate integration of open and proprietary data - no data marketplace yet

Considering the datasets that concern cities and liveability, a number of platforms exist that facilitate the collection and distribution of data. The Rotterdam Open Data portal collects, prepares and distributes all the available open data from the municipality. Furthermore, it engages with a community of business and app developers to stimulate innovation. There are also commercial platforms that collect, prepare and (re)package open data, like esri, Factual and Sense-OS. These platforms enable their customers to combine open datasets and data acquired from, for instance, TomTom, with their own proprietary datasets. Sense-OS focuses on sensor data, and provides a platform to combine several datasets and, like esri, analytical services that enable customers to build and host apps. A platform that also focuses on open data, but specifically on data generated by the Internet of things, is UK-based Thingful. Thingful positions itself as a signpost for ‘the public internet of things’ and aspires to foster innovation around public IoT datasets and the sensors and devices that generate them. Thingful does not collect and store that data itself, but rather makes the data findable and provides interaction between its users to create (local) communities around them.

But these examples mostly focus on open data or the combination between open data and the proprietary data from their clients, and possible some data from established (re)sellers like TomTom. This study found no platform that brokers proprietary data from the city on a bigger scale between all kinds of players.

Existing information intermediaries like Experian, Factual and GfK do collect proprietary data to provide market intelligence for their clients, but they do not facilitate direct interaction between various players. Occasionally companies acquire datasets from other organisations (TomTom for instance, acquired for some time data from telecom provider Vodafone to develop its real-time traffic information) but this is not common practice. Current interactions between open data and proprietary data in the city are depicted in Figure 5.

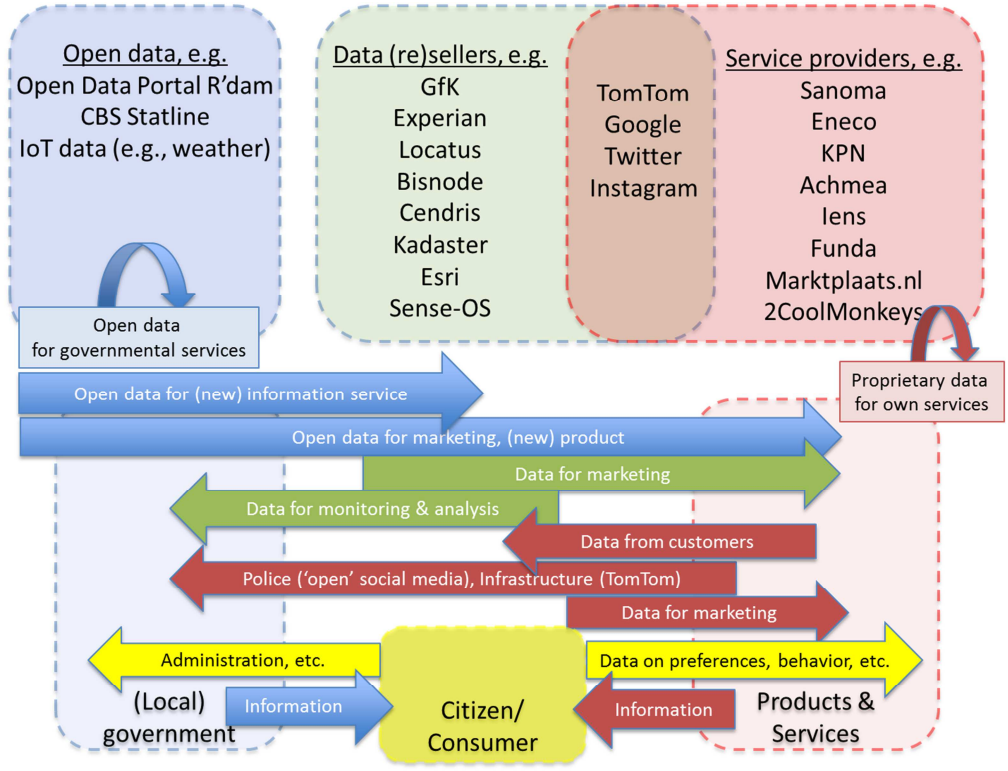


Figure 5: Interaction between open and proprietary data

4.5 Combining global tools with local data

A consideration for players who use data for their services is what kinds of tools are required and to what extent off-the-shelf tools suffice or when are tailored or even in-house built solutions necessary. In many instances off-the-shelf business intelligence and data analytics tools are able to perform the required tasks. The most common providers are IBM, SAP, SAS, Qlikview, Google Analytics and, when dealing with geo-data, esri and Google Maps, as well as a myriad of tools for specific tasks such as e-mail marketing or mouse-clicks. Google, esri and Sense-OS not only provide data and analytics services but also provide tools for developers to build new apps. esri, for instance, works with 2CoolMonkeys who build open data tools on top of esri platforms, for use by government agencies. This construction allows 2CoolMonkeys to profit from the existing infrastructure from esri, while strengthening the relationship between esri and its governmental clients. Funda and Iens deploy Google Maps for their services.

Eneco, Sanoma, Funda and Iens indicate that for tasks that are very specific to the organisation generic tools are often not suitable. For strategic use of data organisations also prefer to perform the analysis in-house. Because Funda expects the data that it captures and analyses to be crucial for the future of its product, it also wants to be able to create and leverage the algorithms that drive its service. For this it is partnering with a third party that is helping them to build this know-how. Organisations also collaborate with knowledge institutes such as universities on advanced use of data tools in profiling, text analysis or predictive analysis. Furthermore, in addition to ensuring data skills when the complexity of analytics increases, organisations such as Eneco, Achmea, Marktplaats and Sanoma are also building an infrastructure that facilitates the democratisation of data analytics. For instance by building 'sand boxes' that enable non-experts to experiment with specific queries or by implementing visualisation techniques.

While most tools deployed by internet services for storage, computing, analysis and visualisation are 'global', liveability is very local and physical. This means that local interventions regarding safety and social cohesion or services that focus on the very local and physical (like restaurants) require local data and knowledge that match the local needs and idiosyncrasies of the local community. The safety in neighbourhood cannot by addressed by centralized data aggregation and analytics by analysts who have no comprehension of the local situation. Somehow, these insights need to be combined with the knowledge of police officers who are active in the neighbourhood – they have to be able to use and interpret the relevant information. Iens described a similar interaction between analytics and the local community: the information that Iens provides to its users is the result from active engagement of its local community. According to Iens, these very local 'intelligence' services cannot be provided by global players that solely work with global, generic data. She likened global players like Tripadvisor to the 'tourist trap' of the internet because its intelligence is not developed by the local community.

Thingful described how their local focus of their search interface stems from their experience that the IoT and its applications are very local. Thingful aspires to facilitate communities find new solutions that match bottom-up local needs. However, as physical products and infrastructures (such as energy grids) become platforms for new services, they might become subject to network effects. The number of users makes the platform more attractive for developers and service providers and the availability of services on top of the device or platform makes it more attractive of users. Wearables, machines and vehicles and even energy infrastructures gain a competitive edge when other companies build services on top of them. Still, a global Internet of Things is not yet here and many 'smart devices' are not connected to the world around them.

4.6 Commodification of the infrastructure

Above we have described how organisations that consider data and data analytics to be one of their (potential) differentiating qualities aspire to acquire in-house knowledge, expertise and data assets. Generic off-the-shelf analytical solutions will no longer suffice and (additional) proprietary solutions are preferred. In contrast, the data infrastructure in terms of storage or computing power is seldom considered to be an important and competitive asset by the interviewees. It is described as a commodity that enables them to perform the differentiating analytics. This kind of

infrastructure and the choices for ‘cloud solutions’ versus ‘local solutions’ and the exact locations of the servers, does become important when handling very large data sets that need to be processed in (near) real-time when the computational power could impact the performance, for instance to instantly provide search results or navigation services. However, for many service providers, this is not considered to be their most differentiating quality but rather a necessity.

4.7 Battle to control the interfaces of the city

In terms of the collection of data and the presentation of data analytics, smartphones and the apps that run on them play a very important role regarding liveability and data in the city, they seem to be one of the most dominant interfaces. Especially for organisations that work with employees who are active on specific locations (e.g. Stadsbeheer and the police) or organisations that rely to some extent on the input from citizens/consumers, smartphones are important tools for data-collection, but also for the presentation of relevant information. This domain is dominated by Apple and Google in terms of hardware (especially Apple), operating system (iOS and Android) and their app stores. Many service providers rely on this infrastructure to deliver their apps. Only few organisations, like TomTom, have built their own hardware. These interfaces, both as a means to acquire data and to distribute output, are strategic strongholds in the data ecosystem because they manage the interaction with the end-users. This means that they are important gateways for any other service provider that wants to interact with users or citizens. Consequently, they function as horizontal platforms that thrive on network-effects.

With the advent of the Internet of Things, the number of interfaces is expected to increase which could signal an intensified battle for the interface: more ‘wearables’ (e.g. TomTom’s Smart Watch, Google Glass and the highly anticipated Apple iWatch), connected devices and smart cars are entering the market. The latter was mentioned by several interviewees as a platform for contextualized services. Furthermore smart TVs, smart meters, and smart homes were mentioned. Especially the smart home could be valuable in terms of independent living for elderly people, which is directly linked to liveability. Google and to a lesser extent TomTom (which has developed a smart watch) are very active in the domain of IoT. Google recently acquired several robotics companies and Nest, a smart meter and home security system, which signifies – together with its Android for the home operating system, a first step in the connected home. Also Microsoft and Samsung are active in the smart home market, as well as organisations that originate in the field of home automation. Energy companies are another type of player that are active in the home with smart meters.

However, although more and more sensors, devices and machines are getting connected it is too soon to see an actual global (or even local) IoT where data is collected, analysed and shared. According to one of the interviewees the current state of the IoT is similar to the early days of the internet before the introduction of commercial web browsers when people were not aware of what was going one beyond the local community they were active in. Still, an emerging ‘smart’ and connected world provides a new battleground and a number of players will want to control these interfaces and transform them into horizontal platforms that will connect users and third party developers.

4.8 Authentication as a service

Several service providers allow users to log-in to their service using their Facebook and/or Twitter username-password combination. This increases the ease of use as users do not have to create yet another username and password, and it also allows the service to leverage the social network, for instance when users more easily post or link activities on Twitter and Facebook (such as an article or a visit to a restaurant). Furthermore, this also functions to increase quality of participation. Iens, for instance, requires users to connect their account to Facebook if they want to write a review to increase the likelihood of authentic contributions. Currently, authentication is merely applied to authenticate users. But as platforms for data sharing develop, including both humans and connected devices, the assertion, authentication and authorization of users and providers of datasets will be increasingly important to facilitate the deployment and integration of various datasets.

4.9 The supporting roles in the data ecosystem

The data ecosystem does not only comprise actors who collect, use and share or sell data and information. As will be described in more detail below, it also includes policymakers and regulators who provide and enforce a regulatory framework, and players such as investors, media and user/citizen communities. These players may not be primary actors in the sense that they provide data services or deploy data in the domain of liveability. However, they do provide the context in which the primary actors operate.

4.9.1 Policymakers and regulators

Policymakers and regulators play an important role in this ecosystem, both in restricting and promoting activities. Firstly, data protection legislation – designed to protect the privacy of citizens – impacts the use of personal data and the combination of datasets that could lead to information that can be used identify citizens. In Europe, and the Netherlands in particular, the use of personal data is subject to purpose limitation, which means that data about individuals cannot be used for new purposes without the consent of that specific individual. However, data protection legislation is organised differently in Europe and in the United States. According to several interviewees this means that there is no level-playing field in the digital services market where companies from different countries have to deal with different sets of rules.

The European Public Service Directive[50] inspired the development of the Rotterdam Open Data portal and other open data initiatives and funds several schemes that aim to foster a European data ecosystem. Another, more indirect example of how policymakers can influence the use of data analytics is the implementation of differential taxes in the Dutch municipality of Deventer for waste disposal: households with more waste pay more taxes. This spurred the development of a data infrastructure where several waste management organisations share data resources to provide this information to the local tax office. It also stimulated the use of data to improve internal operational processes of the waste disposal

[50] http://ec.europa.eu/digital-agenda/en/european-legislation-reuse-public-sector-information

organisations, such as a more optimal deployment of resources such as employees and trucks.

The government can also provide subsidies for SMEs and start-ups. Peerby, for instance, received some initial funding from AgentschapNL.

4.9.2 Accelerators and investors

In addition to platforms that facilitate interaction on the data level, there are also accelerator platforms that facilitate the interaction between technical start-ups, mentors and investors, like Startupbootcamp, F6S and Rockstart (which supports local sharing-service Peerby), or in-house accelerators such as Sanoma (also an investor in Peerby) which spun out Hubly, a Dutch marketplace for freelance content creators. Startupbootcamp has initiated programs in various European Cities, each with a different focus, such as mobile, NFC, health and, in the near future, big data. These programs have several sponsors, such as Amazon Web Services, and connect the start-ups to mentors from the industry. For the programme on big data Startupbootcamp is currently in contact with companies like IBM, Cisco and Intel. For these companies, the accelerator programs are an opportunity to connect with innovative people outside their own organisation. It enables them to forge new strategic partnerships to strengthen their position in the market, invest in (or acquire new companies) and business development in general.

4.9.3 (Local) media

(Local) media have shown interest in the information that service providers collect on cities. Marktplaats and Iens, for instance, are occasionally approached by media with a request for lists regarding the most active cities regarding barter (Marktplaats) or trends regarding restaurants and bars. However, this has not resulted in structural adoption of new kinds of analytics in these organisations. It seems that media have a more important role in the uptake of new data services, especially when these services are directed at consumers. Peerby noted that media attention from major national television networks and print media stimulated the uptake of its service.

4.9.4 (Local) communities

Some service providers rely heavily on the participation of their users, both for their active contributions (e.g. Iens posts reviews) and passive contributions (e.g. TomTom automatically collects data via their users on the road – although their users can also actively contribute). However, this kind of community involvement only concerns the generation of data. Other kinds of community involvement concern direct interaction and are not data intensive although they do provide value to the service provider (e.g., the police and Buurt Bestuurt) or the involvement of local users to realise part of the service (e.g. Marktplaats.nl and Peerby) or to foster new innovations (Thingful regarding the IoT).

4.10 The data-ecosystem of liveability in Rotterdam: many vertical services and some horizontal platforms

Figure 6 depicts the various types of products and services related to the value creation process of data regarding liveability in Rotterdam; from the interfaces that generate data to the interfaces that provide the output, and the various steps in between: data capturing, authentication, preparation, integration, storage and

analysis. Also several different output can be distinguished based on the output generated based on the data, such as visualization, action, and the sharing or selling of data. Additionally, it also includes the various roles that shape the overall context of the data ecosystem, such as research, regulation, venture capital, media and user communities as described in the previous paragraph. This figure aims to captures the data ecosystem of the liveability domain in the city of Rotterdam.

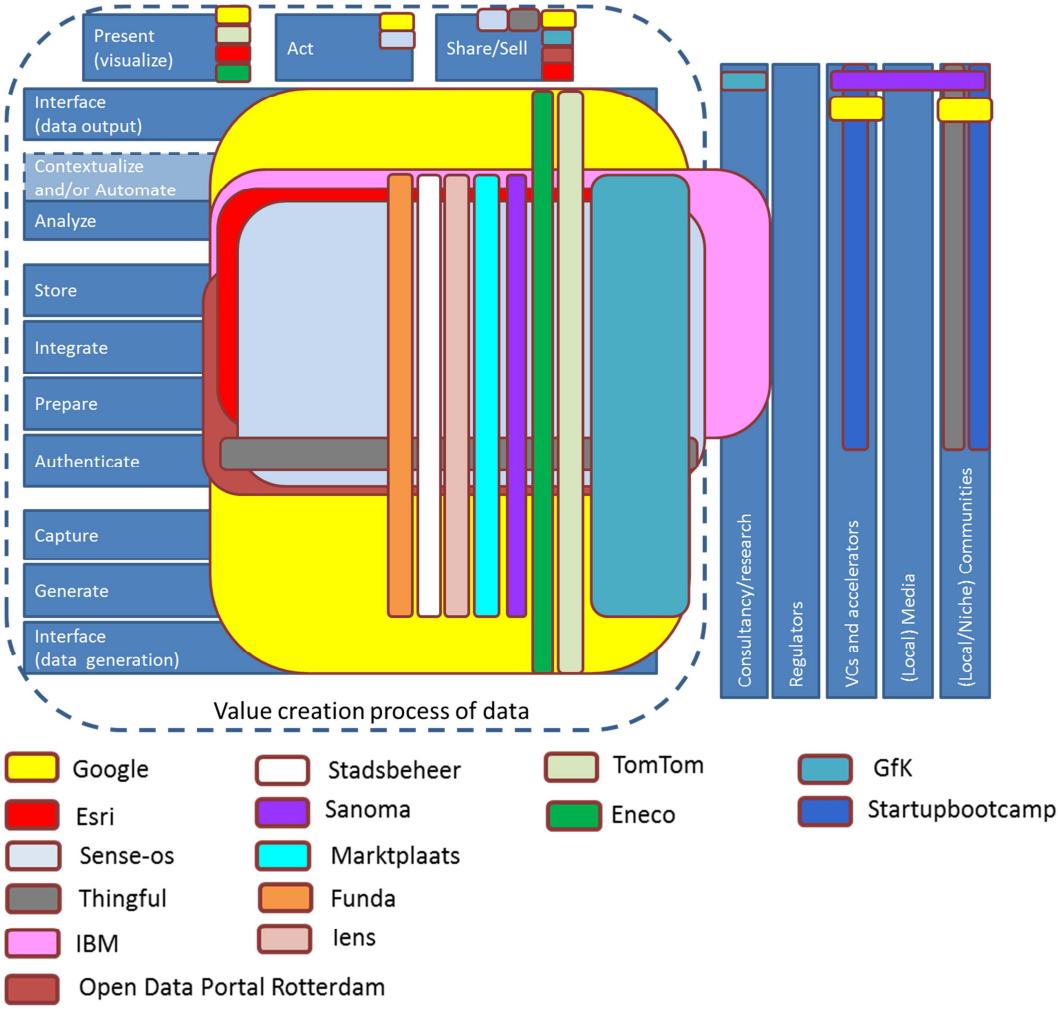


Figure 6: The data ecosystem of liveability in Rotterdam

Figure 6 also includes some of the organisations from domain of liveability, plotted on the various roles in the value creation processes that drive realize data-driven innovations in the domain of liveability in Rotterdam. This figure does not include all organisations that were interviewed, but a selection of organisations to illustrate the current dynamics regarding the fulfilment of the various roles in the data ecosystem. Note that an organisation can fulfil multiple roles in the data ecosystem.

The data ecosystem of liveability in Rotterdam appears to comprise many vertical services – both governmental and commercial, each focused on a specific domain such as housing, restaurants, the public sphere and energy – and a few horizontally oriented platforms that provide a specific role (or number of roles) in the value

creation process for many different users, often B2B. These horizontal platforms are not only concerned with the most generic aspects of the data value creation process such as storage or processing, but mainly analytics (e.g., GIS-systems from esri and Google, or other kinds of analytics (e.g., IBM and sense-OS) that can be used by other, more vertically oriented developers to build new services) or data-platforms that facilitate the finding, sharing and distribution of (often Open) data, such as the Open Data Portal Rotterdam, esri, Thingful and Sense-os). As described above, there are several data platforms, but this study did not find a marketplace that enables (direct) interaction between organisations to share or sell data with/to other parties. As described above, in section 4.4, the data that is being collected by commercial service providers is currently only sparsely integrated in the governmental domain.

Whereas the vertically oriented services in the domain of liveability seem more directly related to the national, the local or very local, and thus liveability, the horizontal platforms have a broader, more international scope. Although players can fulfil multiple roles in the data-ecosystem, most of the organisations in this study either focus on a specific vertical or (aspire to) function as a horizontal platform. The most notable exception is Google – covering almost all aspects of the data ecosystem. But also IBM is complementing its more generic offering with more vertically oriented analytical tools, although it is not active in the 'bottom' three roles in the value creation process like Google.

As described above, the advent of the Internet of Things fosters the development of new interfaces in addition to PCs and mobile phones: for instance smart cars, devices and wearables. Still, only TomTom (with their proprietary hardware and smart watch), Eneco (with their smart meters) and Google (with Android, the Google phone and more recently Nest) have positioned themselves as direct interfaces. However, this could change in the future as more and more objects become connected and 'smart'.

If we look at the interaction between incumbents and new players, several observations can be made. Existing players in the domain of liveability are often either vertically oriented players (like the police, Stadsbeheer or Funda) or traditional IT-providers like IBM that have a strong position in horizontal IT and BI markets. The vertical players often deploy data analytics to improve their existing value propositions. This study found that currently the vertically oriented incumbents only sparsely provide truly new value propositions and when they do the value proposition has a vertical bent that suits their core domain (e.g., Iens' reservation system for restaurants). Traditional IT and BI vendors try to improve and extend their offering to meet and increase demand for their products in both existing and new markets.

New players seem to look for both vertical and horizontal niches. As described earlier, startups that focus on elements of liveability and/or 'the city' create new services deploying open data and/or data acquired from their users via proven interfaces such as PCs and smartphones. As the Internet of Things proliferates, these services could also emerge around new interfaces such as cars, devices and wearables. Also new analytics providers seem to have a more vertical focus, often building on top of existing, generic analytical platforms. The incumbent horizontal players try to strengthen their position in vertical market by developing new vertical

services, or acquiring the most successful vertically oriented startups in promising markets. These new analytical services enable other organizations the realize data-driven innovations (e.g. the use of Twitcident by the police).

There are also horizontal niches, as the emergence of platforms like Sense-OS and Thingful illustrate. Both are specifically focused on sensor-data which is still in its infancy. These platforms allow other players to build new services on top of them.

5 Conclusion

Smart cities try to leverage data to align and integrate infrastructure, planning and management, and human services as a system of systems – with the goal of making cities more desirable, liveable, sustainable, and green. In the future, these cities are expected to have 'interfaces' – ecosystems of data layers, of platforms, services and apps that access (open and proprietary) data on the city that enable meaningful analyses for policy making and the development of commercial services.

This study investigates whether such a data ecosystem emerges in the domain of liveability in the city of Rotterdam. It consists of a network analysis of the domain of liveability in the city and of the data landscape to determine which organisations deliver services for liveability in the city, which organisations collaborate and create new services and platforms, an analysis of the data that is being used and the value propositions that emerge, and an investigation of the main driving and impeding forces of this development. Based on these analyses, the previous chapter described the emergence of a data ecosystem in the domain of liveability in the city. Furthermore, this final chapter aims to answer the remaining question of how the emergence of this data ecosystem impacts the domain of liveability. Furthermore, it presents some recommendations for further research.

5.1 (Semi-)public organisations retain control over the liveability domain

Services in the domain of liveability are delivered by governmental, semi-public and commercial organisations. Furthermore, government organisations perform long-term focused activities, such as policy making and urban development. These activities seem to rely on their existing infrastructure of knowledge building and the acquisition of insights by using monitoring tools devised in collaboration with research organisations and public/private organisations. Still, some of these monitoring tools do deploy data from third parties, such as CBS, Kadaster and commercial information brokers like Experian, Cendris and Bisnode and integrate this into their 'knowledge infrastructure'. Due to their long-term focus, these activities do not process real-time data for real-time decision-making. They look for trends and developments over a longer period of time, profiling neighbourhoods to devise or adjust policies. Rotterdam has integrated various monitoring tools such as the Safety Index, the Physical Index and the Social Index to create an integrated overview of a neighbourhood, which could support integrated interventions. This notion of integration is also at the heart of the Early Warning System in Almere in which the goal is to overcome the siloed structure of governmental services.

The deployment of data and data analytics for daily governmental operations has a different dynamic. Data and data analytics play, for instance, a crucial role in the maintenance of the public sphere and for safety purposes. Various departments of Stadsbeheer as well as the police process data that is collected by themselves, as well as data from citizens (e.g. via the *Beter Buiten* app, Police app) and – in case of the police – social media and other open online sources that are analysed with tools such as *Twitcident* and Costoo. In a few cases the insights gathered are contextualised, especially in the case of safety, where the policy apply the concept of predictive policing and work with profiling techniques of neighbourhoods. But

overall these kind of analyses are only beginning to take shape. Regarding social wellbeing several pilot projects are undertaken, such as Almere's Early Warning System. Automated decision-making based on data analytics was not found to be deployed yet.

The impact of data analytics on liveability via commercial services appears limited. There are services that have a direct link to liveability, by collecting data (such as the *Beter Buiten* app), deploying data where the analytics supports business operations (Achmea) or that are related to liveability via their services (such as the *Bomenspotter* app and Peerby). These services do not (yet) deploy data analytics for contextualisation. There are many more services that collect data about citizens and the city. But because of their focus on their existing business model, which is a crucial factor in the strategy of deploying data and data analytics, the impact of such services is expected to be limited in the domain of liveability. Of course, this could change. As Startupbootcamp mentioned, even though most (new) services may not primarily be focused on data as their core business, scaling up means, eventually, that data will be part of the service, especially when the service has an online component. But this could also apply to physical goods and services that become subject to 'datafication', for instance when devices become equipped with sensors and connectivity. Although the impact of commercial service providers may be limited at this point, the impact of data and data analytics will eventually increase when new services arise that have a direct relation to elements of liveability.

5.2 Data sources on liveability: from G2C to G2B2C to G2B2T2C2All

If we take a different focus and look at relevant data on liveability and the city and at the actors who collect these data, there does appear to be a shift in control from the government to commercial service providers and citizens/consumers. Many of these service providers might not have direct impact on liveability, at least for now, but they do collect data about cities and citizens in all kinds of domains. This may mean that in the future the difference in control over these data may be observed within those domains.

Traditionally, relevant data regarding liveability is collected by government services and semi-public organisations, such as health insurance companies and housing corporations. However, due to the datafication of society, more and more fine-grained data about the city that could be of use for the domain of liveability – especially in (nearly) real-time, is collected via (commercial) products and services. Google collects data about search and browsing behaviour, Facebook and Twitter collect data on our social interactions and our locations. More locally oriented service providers such as Iens, Funda, Peerby and Marktplaats collect data on where we live, eat, shop and share. Utility providers such as energy companies detect anomalies in our consumption patterns that may indicate potential problems such as fraud, while telecom providers are able to detect aggregated patterns on mobility in cities.

But the data explosion is not only a matter of companies collecting data about our behaviour. In addition to these service providers, citizens are also starting to collect and exploit their own data. Initially via apps and online services, but this could expand via wearables - like smart watches, Google glass - and smart vehicles, smart homes and smart devices. The picture that is being painted about our cities,

citizens, ‘things’ and the way they interact with each other becomes much more fine grained. We are shifting from G2C to G2B2C and possibly in the near future to ‘things’ as well: G2B2T2C. And citizens may contribute as well, which may even lead to something like C2B2T2C2All. However, although the data that is collected about cities increases, this does not mean that this data is also being deployed in order to maintain and improve liveability as it is being controlled by commercial services and in some cases by citizens.

Overall, this study found only limited spill-over from data collected by commercial service providers regarding cities and their citizens. Exceptions are social media data and, in some cases, data from TomTom to improve the physical infrastructure. Currently, there seem to be no mechanisms in place that facilitate a further exploration of these new datasets in order to improve liveability. Commercial services are mostly focused on the deployment of data for their primary business processes rather than exploring potential use for third parties, like governmental services or totally new services that could impact liveability. Other than the traditional information brokers that aggregate and sell market intelligence like Experian and GfK, there appears to be no platform that is used in this domain to facilitate sharing, selling and acquisition of proprietary data. There is no data market yet that fosters the exchange of proprietary data. And as long as the core of these services and their underlying businesses models are not directly linked to elements of liveability it is not likely that exploitation of all these data has an impact in this field.

5.3 Looking ahead

This study found that even though many data initiatives emerge, the domain of liveability in the city is still most strongly impacted by governmental services. (Semi-)public organisations appear to control the domain and its main policies and services for now. Regarding the possession of data, however, commercial organisations appear to acquire more and more data that is usable for this domain. A question that arises is, thus, when data integration becomes increasingly important, whether the domain will retain its largely public character or that new commercial services will take over (parts of) the public tasks and functions. Especially since also the interfaces that collect the data and facilitate and shape the interaction between services and end users, such as mobile phones but also smart cars, smart meters and wearables, are increasingly owned by commercial organisations. As long as government is successful in its realization of liveability the incentive for companies to become active players in this field may be limited. But this may change, especially when citizens are expected to become more active participants and have to rely increasingly on self-organisation.

What could these changes mean for liveability? Businesses would play an important role to facilitate citizens’ participation. And when business models from existing players or start-ups collide with elements of liveability, the current balance between public and private may change, including the impact of their deployment of data and data analytics. There are examples of companies that are already active in certain fields of liveability, such as services like Peerby and Tinder that mediate interaction between neighbours and singles respectively, which could impact social cohesion. Or Google’s Nest, which acquired home surveillance company Dropcam, may impact the domain of physical safety. This is already visible in the US, where the

public sector is traditionally smaller in size, and many apps have emerged (apps that map potholes in cities and thereby influence operations are a notable example) that perform this function.

The interaction between public and private players may also be exacerbated when the current focus on transactional and geo-location data shifts to the integration with social media data and sensor data, which are more often owned by the private sector and/or citizens. Another impulse for increased interaction between governmental and private services could be provided by (smart and pro-active) marketplaces that facilitate exploration of external datasets from third parties via direct interaction, rather than aggregate information services from current intermediaries. However, it might be expected that, as with the deployment of data analytics, these kind of marketplaces follow existing interest in data sharing and selling, and thus potential revenue, rather than the other way around.

And it remains to be seen what may happen to the domain of liveability and how government may or may not adapt its policies. In other sectors and markets, the emergence of ‘killer apps’ has had great impact, such as AirBnB in the hotel branch and Uber in the taxi branch which rely heavily on data analytics. In these markets the ‘interface’ to the market has led to a network of competitors (with very low thresholds for entry), threatening the existence of traditional service providers and possibly changing these sectors forever. The question is whether governments are willing to let that happen in the domain of liveability.

To our knowledge, this study represents one of the first empirical studies of big data and smart cities, actually mapping actor-networks and presenting findings on how the interface of the city is likely to emerge and what are the main implications. As often, developments and impacts are different than what is expected. Also, many technological developments that are expected to have mainly positive impact, often also have (expected and unexpected) less desirable consequences. As these developments may impact all citizens, further research is recommended to understand the ongoing developments in the field of smart cities.

ANNEX 1: SHORTLIST OF ACTORS/ORGANISATIONS

Organisation	**Name interviewee**
Early Warning System Almere	Mr. Steeg, T. van der
Iens	Mrs. Boswijk, I.
Leefbarometer (Liveability Index)	Mr. Kleefstra, H.
The public matters Author of 'De stad als interface'	Mr. Waal, M. de
Open Data Rotterdam	Mr. Groot, F. de
Peerby	Mr. Weddepohl, D.
Rotterdam Stadsbeheer (Public Sphere)	Mrs. Goederen, K. de
TNO – Quality of Life	Mr. Egter Wisserkerke, N. van

ANNEX 2: INTERVIEW PROTOCOL

For the interviews initially two categories of interview protocols were used for 1) organisations that use data and data analytics to support their operations and products, and 2) organisations that provide data-services (data sources, analytics, storage, etc.). The interview protocol covered all relevant aspects for answering the research questions of this study. The duration of the interviews varied between 30 and 60 minutes. Most interviews were recorded and were conducted one of the authors.

Interview 'users' of data and data analytics

1 Please provide a short description of your company and your role within the organisation.
2 To what extent are you systematically using data analytics? For which purpose/goal/end?
3 What were the most important drivers to start using data analytics?
4 Please provide a description of the process that led to the decision the apply and implement data analytics within your organisation.
5 What are the key technologies and tools you are using?
6 What have been the most important barriers before you could systematically use data analytics?
7 What are the most important barriers you are faced with when using data analytics?
8 Could you give us some specific examples of big data analytics that you are using? In which departments or part of your organisations are they using it? What is the output of the analysis and how is this being used?
9 What is or has been the actual result of this big data application (e.g., efficiency, new products, ...)
10 What types of data are you using for these analyses (e.g., historical/real-time, structured/unstructured, human/social/sensor)
11 What is the source of these data? Generated in-house of retrieved from third parties?
12 If third parties are the source for data, how is this data payed for? (e.g., licenced?)
13 Which steps in the value creation process do you fulfil yourself and for which steps do you rely on or collaborate with third parties (for transport, storage, management, analysis, visualization, security, etc.)
14 Regarding parties you collaborate with: which parties? And why these parties? Did you consider other parties as well?
15 What are your short-term and long-term ambitions for your organisation and how does data analytics play are role?
16 Do you also produce data products or services (e.g., analysis of own data, consultancy) as part of your portfolio?

Interview 'provider' of data services

1 Please provide a short description of your company and your role within the organisation.
2 Which types of products or services do you deliver within this process?

3 What are the key technologies and tools you are using?
4 Are you using open standards? If so, under which conditions are they being used or can they be used?
5 Do you collaborate with other parties/stakeholders to provide the service/product you are delivering? What is their complementary value?
6 Are there other companies that offer the same kinds of product(s)/service(s) [name company]? If so, what are the most important competitors?
7 How does [company] differentiate itself from competing products/services?
8 What kinds of customers does [company] cater to with [product]? For what kinds of purposes do customers use the product/service? What are their drivers?
9 What are the most important barriers that customers of a product/service like [product] have to overcome (financially, organisational, technological, legal, …)
10 What are the most important ambitions from [company] regarding data services in general and [product]-like products/services specifically?

ANNEX 3: FULL LIST OF INTERVIEWEES

Organisation		**Name interviewee**
1	2CoolMonkeys	Mr. Olst, R. van
2	Achmea Health	Mr. Egberts, B.
3	Buurt Bestuurt	Mrs. Eykelen, C. van
4	Early Warning System Almere	Mr. Steeg, T. van der
5	EDM	Mr. Feltzer,
6	Eneco	Mr. Groen, T. de
7	Funda	Mr. Peeperkorn, K. Mr. Borms, T.
8	Esri	Mr. Vermeij, B.
9	Google	Mr. Versteeg, D.
10	IBM	Mr. Blommaart, J.
11	Hogeschool Rotterdam	Mrs. Lemmens, J.
12	Iens	Mrs. Boswijk, I.
13	Leefbarometer (Livability Index)	Mr. Kleefstra, H.
14	KPN	Mr. Reichel, T.
15	Marktplaats	Mr. Mooy, R.
16	National Police	Mr. De Hengst , S.
17	Open Data Rotterdam	Mr. Groot, F. de
18	OMA / AMO	Mrs. Baird, L.
19	Peerby	Mr. Weddepohl, D.
20	Plan B	Mrs. Boersma, E.
21	Rotterdam Stadsbeheer (Public Sphere)	Mrs. Goederen, K. de
22	Rotterdam Stadsbeheer (Safety)	Mrs. Hira, D.
23	Sanoma	Mr. Kieft, S.
24	Sense-OS	Mr. Larsen, J.
25	SOR	Mr. Pollen, S.
26	Startupbootcamp	Mr. Shannon, A.
27	The public matters	Mr. Waal, M. de
28	TNO – Quality of Life	Mr. Egter Wisserkerke, N. van
29	TNO – Information Society Author of ‘Social Media DNA’	Mr. Vries, A. de
30	TomTom	Mr. Hania, S.
31	Umbrellium (Thingful)	Mr. Haque, U.
32	Veiligheidsindex (Safety Index)	Mrs. Berge, S. van den
33	Veldacademie	Mrs. Hoppner, R.

ANNEX 4: EXAMPLE OF THE NETWORK ANALYSIS

Text from Box 1 on smart cities

The EU defines the smart city as “a system of people interacting with and using flows of energy, materials, services and financing to catalyse sustainable economic development, resilience, and high quality of life; these flows and interactions become smart through making strategic use of information and communication infrastructure and services in a process of transparent urban planning and management that is responsive to the social and economic needs of society”.

Example of relations between actors (players, technologies, concepts) from the highlighted text above. These actors and their relations would be used as input for the Citoscape analysis.

EU	Smart City
Smart City	System
Smart City	People
Smart City	Energy
Smart City	Materials
Smart City	Services
Smart City	Financing
Smart City	Sustainable economic development
Smart City	High quality of life
Smart City	Resilience
System	People
System	Energy
System	Materials
System	Services
System	Financing
People	Sustainable economic development
People	High quality of life
People	Resilience
People	Energy
People	Materials
People	Services
People	Financing
Sustainable economic development	Energy
Sustainable economic development	Materials
Sustainable economic development	Services
Sustainable economic development	Financing
Resilience	Energy
Resilience	Materials
Resilience	Services
Resilience	Financing
High quality of life	Energy

High quality of life	Materials
High quality of life	Services
High quality of life	Financing
Smart City	Smart
System	Smart
Smart City	Information
Smart City	Communication infrastructure
Smart City	Transparency
Smart City	Urban planning
Transparency	Urban planning
Smart City	Strategic use
Strategic use	Information
Strategic use	Communication technology
Smart	Information
Smart	Communication technology
Smart	Services
Urban planning	Information
Urban planning	Communication infrastructure
Smart City	Management
Information	Management
Communication technology	Management
Smart City	Responsive
Smart	Responsive
Management	Responsive
Responsive	Social needs
Responsive	Economic needs
Smart City	Society
Management	Social needs
Management	Economic needs
Social needs	Society
Economic needs	Society

ANNEX 5: DATA VALUE CREATING MECHANISMS

This table comprises the value creation mechanisms of data analytics that were mentioned during the interviews. It distinguishes the deployment of data analytics to enhance existing propositions and the deployment of data analytics to create new value propositions. It also distinguishes current practices or ambitions/expectations for the future. Note that the number 'current practices' and 'future' do not necessarily add up to the 'number of interviews'; in some cases organisations are currently deploying data for a certain goal, while expecting this to become more advanced in the future.

Enhancing existing propositions	Number of interviews that mentioned the VP	Current practice	Future	**New value propositions**	Number of interviews that mentioned the VP	Current practice	Future
More fine-grained and integrated insights regarding customer base and user behaviour (segmentation and profiling)	8	7	3	Tools that can perform data analytics (e.g., for BI, advertising, site performance, sentiment analysis)	20	20	10
Improving marketing strategies by monitoring their impact (data-base driven marketing)	6	6	-	Personalised, adaptive and predictive services for consumers (e.g., health, energy consumption, driving behaviour)	10	4	6
Support policy design regarding liveability	6	6	-	Tools that facilitate citizen/user participation and/or interaction	8	7	2
Monitoring impact policy regarding liveability	5	5	-	(Open) data platforms for data distribution	7	7	-
Monitoring product performance (e.g., site performance)	5	5	-	Platforms that enable data-search, sharing and development of new services with available data on top of the platform	7	4	3

Optimising supply and demand (using predictive analytics and profiling techniques)	5	5	3	Smart devices in the home (e.g., smart thermostats)	6	3	3
Adapt communication strategy to customer profiles	4	2	3	Smart cars that are connected and can collect, process and share data and support information services	6	-	6
Providing insights that enable more integrated interventions regarding livability and safety in particular in neighbourhoods	4	4	1	Consultancy regarding advanced data analytics, technology platforms, implementation and management	6	6	-
Monitoring the state of (objects in) the public sphere	4	3	1	New data-based services regarding the public sphere (e.g., Bomenspotter)	5	5	-
Optimising online user experience (e.g., via A/B-testing)	3	3	-	Providing maps and GIS-systems as a service to visualise geo-data and information	5	5	-
Providing more fine grained insights regarding safety in neighborhoods for operational services, including historic developments	3	3	-	Authentication (e.g., using social media accounts)	4	2	2
Allocation of resources (e.g., man hours and vehicles)	3	3	-	Software as a service, based on existing data-infrastructure (e.g., fleet management)	4	2	2
Optimising logistical processes	3	3	1	Benchmarking and market intelligence (e.g., restaurants, houses and	3	1	2

				health care providers)			
Predictive interventions regarding liveability (e.g., police)	2	2	2	Information services regarding the use and state of public infrastructure (e.g., roads)	3	2	1
Optimize customer service based on user profiles and history	2	2	-	Visualization tools to support data analytics and -use	3	3	2
Accountability in terms of policy and SLAs	2	2	-	Smart wearables (smart watch, Google glass)	3	1	2
Inventory management	2	1	1	Smart infrastructure (e.g., roads, energy grid)	2		2
Fraud detection	2	1	1	HPC Infrastructure as a service	2	1	1
Information based on analytics for PR purposes	2	2	-	Early warning for natural events (fire, earth quake)	2	1	1
Finance and credit management	2	2	-	New propositions based on data analysis for marketing purposes	1	1	-
(Semi-) automated processing user contributions (based on text-analysis)	2	2	1	SDK for the city leveraging open data	1	1	-
New distribution channels for information and services	2	-	2	Proprietary data for scientific purposes	1	1	-
Supply chain optimisation	1	-	1	Customised product/service delivery	1	-	1