

Uncovering the purchasing of private sector data by governments: State of play in the Netherlands

Iryna Susha^{a}, Sofie de Wilde de Ligny^{a,b}, Mirko Tobias Schäfer^b, Fredo Schotanus^c*

^a Utrecht University, Copernicus Institute of Sustainable Development, Princetonlaan 8a, 3584 CB, Utrecht, The Netherlands, i.susha@uu.nl

^b Utrecht University Data School, Princetonlaan 8a, 3584 CB, Utrecht, The Netherlands

^c Utrecht University School of Economics, Princetonlaan 8a, 3584 CB, Utrecht, The Netherlands

* Corresponding author

Abstract. To address complex societal challenges, governments increasingly need to make evidence-based decisions and require the best available data as input. As much of relevant data is now in the hands of the private sector, governments increasingly resort to purchasing data from private sources. There is, however, scant empirical evidence and a lack of understanding of how government organizations go about data purchasing. This research is the first effort to create an evidence base about what data Dutch governments purchase from whom, how, and what for. In this study, we conducted a mapping to identify cases of Dutch governments purchasing private sector data (and related services) in the context of societal challenges. Our results map the buyers and sellers of data and data services in the Dutch context, as well as the types of data and data services sold, and in which policy do-mains. We further discuss our results in view of the data monetization routes identified from the literature.

Keywords: Data Purchasing, Data Monetization, Business-to-Government Data Sharing, Procurement, Netherlands

1. Introduction

For addressing societal challenges, public management increasingly seeks access to relevant data. Data is expected to provide necessary insights, and support evidence-based decision making and the measurement of policy impact. For instance, to be able to maintain and use infrastructure more efficiently, governments need satellite data. Municipalities need data about, for instance litter, the quality of plants/trees, data about social care, to name a few examples. The problem, however, is that relevant data is no longer solely in the domain of the public sector, but increasingly also in the hands of the private sector (Mayer-Schonberger & Ramey, 2022). Although governments, such as national statistical offices, have long been using private sector data and are often dependent on it, due to the growing 'datafication' and 'platformization', businesses now hold more and more of key relevant data and advanced governance capacities (Ruppert, Isin & Bigo, 2017; Van Dijck, Poell & De Waal, 2018; Sharon, 2020).

To gain access to private sector data potentially valuable for addressing societal challenges, governments can follow several routes. The first route, which has attracted most academic attention, is governments and businesses engaging in voluntary Business-to-Government (B2G) data sharing under the banner of 'data philanthropy' or 'data collaboratives' (Verhulst & Sangokoya, 2015; George, Yan & Leidner, 2020). The second route is governments imposing data sharing obligations on the private sector to share data for public interest purposes in certain situations. And the third route – in focus of this research – is governments purchasing data from the private sector, either through a procurement procedure or alternative means. Exploratory research in the EU finds that data purchasing is apparently the most widely considered route (Micheli, 2022) but it has not yet been systematically researched. At the same time, the market for data is rapidly emerging, including in the B2G domain, and it is characterized by several unique challenges, such as the risk of monopolistic data suppliers and monopolistic prices, high transaction costs and risks of data sharing, and a lack of incentives for private firms to contribute to the production of public benefits (Martens & Duch-Brown, 2020). Moreover, there has been limited research on data-driven decision-making in public procurement (Langseth, 2024).

Therefore, it is important to better understand how governments increasingly interact with data providers. While there is extensive literature on procurement in general and specifically in sectors such as healthcare and construction, this is not the case for data purchasing by governments. There is scant empirical evidence and a lack of understanding of how government organizations go about data purchasing. This is problematic because data purchasing can differ substantially from general procurement or procurement in specific sectors. For instance, data purchasing can be highly intertwined with other parts of the purchase, sometimes to the extent that purchasers might be unaware they are also purchasing data. Other differences relate to the pricing model, ownership, control, privacy, and the accessibility of the data.

There has not yet been any systematic effort to gather evidence about when, under what conditions, and why government organizations choose to purchase private sector data. This motivates our research, hence, in this paper we aim to empirically answer the following research question: *What is the state of play as regards purchasing of data by Dutch governments from the private sector for addressing societal issues?* To answer this question, we conducted an exploratory study to identify 'cases' of Dutch governments purchasing private sector data for addressing societal issues. We further analyzed these cases based on the parties involved, the content of the purchase, the purpose, the procedure used (who, what, how, and what for). The Netherlands presents an interesting case study due to the quite developed nature of its public procurement function and several instances of

governments utilizing data from providers. The insights gained from this country are likely to hold relevance for many others as well.

2. Data purchasing: background and state of the art

As mentioned above, data purchasing is one of the modalities of accessing private sector data, next to for instance mandatory data sharing or data donations. Academic literature on business-to-government data purchasing is very limited and focuses on government agencies at the national level, primarily in the US, and on certain types of data, for instance location data or satellite data (Borowitz, 2019; Shenkman et al., 2022). The study by Micheli (2022) was the first effort to explore data purchasing by government organizations in the EU, as one of the approaches to access private sector data. In a study of 12 European cities, Micheli (2022) found that purchasing of private sector data by municipalities poses several challenges, such as lack of control and risk of introducing inequalities among municipalities. Certain types of data may be in higher demand, such as telecom data, online payment data, sensor data, data from IoT devices (VNG Re-alisatie, 2017). The price of such data tends to be high, as the data owners or license holders know the value of these data and may even split large datasets into smaller and more profitable ones (VNG Realisatie, 2017).

There appears to be a similar void in the literature outside the digital government domain regarding how companies acquire external data. Jarvenpaa & Markus (2020) explicitly call for attention to the concept of *data sourcing* in Information Systems research. Data sourcing is defined as “procuring, licensing, and accessing data (e.g. an ongoing service or one-off project) from internal or external suppliers” (Kotlarski et al., 2018). This term is used in the literature also in the context of data partnerships, and does not necessarily implies paying for data. Furthermore, in the IS literature, the terms purchasing, sourcing, procuring, or buying are sometimes used as synonyms referring to “transaction where a particular good is transferred between two organizations” (Schneider et al., 2013). As rightly pointed out by Krasikov et al. (2022), using external data is not new to companies/organizations and has been on the rise since 1990s, but in the literature this has been associated with simply “getting the data” and any systematic inquiry into how this is being done has been lacking. Our discussion above focused on the buyer perspective of purchasing data which, as we conclude, is vastly under-developed in research. However, if we look at the seller perspective, there is much more that can be learnt from the literature on **data monetization** which has been expanding in recent years. Data monetization goes beyond selling the data directly for money, it also occurs when companies use data to create value driven products, convert data and analytics into financial gain or other tangible benefits (Ofulue & Benyoucef, 2022). Based on Parvinen et al. (2020), we consider three data monetization routes in our research (see Figure 1): (1) selling data; (2) selling analyses; and (3) selling data-based services.

Selling data is granting ownership of raw or aggregated data to the buyer (Parvinen et al., 2020). This route can also concern companies that generate data as part of their normal operations that is then marketed to a third party for a purpose other than that for which it was collected (Schroeder, 2016). This practice can also be referred to as data repurposing (Gunther et al., 2022). An example is telecoms offering call details records for population movement analytics and related new purposes. In the literature, this model is also termed as Data-as-a-service, whereby raw or aggregated data (including as combined from various sources) is delivered (Hunke et al., 2022). This route can also include data intermediaries (e.g. data brokers) who ‘make a living’ from selling data acquired from external parties (Christl & Spiekerman 2016). Ritala et al. (2024) discuss this route as productization of data. This model can also evolve into subscription-based data access, whereby “users subscribe to a data provider for a

specific period and pay for data access based on the selected subscription plan” (Ofulue & Benyoucef, 2022).

Selling data analyses is the second data monetization route, whereby analyses based on data are offered (in the form of insights, interpretations or benchmarks) but access to the original data is restricted. This route usually “includes different consultation, audit, and diagnostic services aimed at analyzing customers’ operations and suggesting potential opportunities for improvement” (Ritala et al., 2024). This embodies the servitization of data (Ritala et al., 2024). Data intermediaries (data brokers, data analytics companies, data marketplaces) are active players using this route as they provide insights by connecting, aggregating, and modelling data from multiple sources (e.g. social media, mobile devices and apps, etc.) and making them available for further uses and users (Beer, 2018 cited in Helmond & van der Vliet, 2023). In this category, we exclude ‘traditional’ consultancy services such as reports, advice and benchmarks, which are widely in use by governments. Instead, we mainly focus on the purchase of (big) data analytics services.

Selling data-based services is the third data monetization route, whereby a new data-based service is created that can provide customers with relevant signals on the environment, help scale how data is delivered using multi-sided business models and can help customers act on insight (Ofulue & Benyoucef, 2022). In this case, companies create new services, usually implemented via dashboards or other digital interfaces, whereby the data comes from the service user (Parvinen et al., 2020). An example is a customer relationship management plat-form that uses customer data. The literature also mentions cases where companies provide clients with means to collect new (object) data (typically sensor-based IoT data) integrated into customer’s workflow (Hunke et al., 2022) as part of a new service. Ritala et al. (2024) discuss this trend as performatization of data where products, services, and data are combined into intelligent and smart solutions that deliver optimized process outcomes to the customer. An example of Google Nest is provided in the literature, whereby Nest offers insights to electricity utility companies on users’ energy consumption (Parvinen et al., 2020).

The three aforesaid routes or combinations of routes are examples of direct data monetization by companies where a commercial transaction occurs for monetary rewards (Ofulue & Benyoucef, 2022). Data monetization also occurs *indirectly* in the case of data bartering (Wixom, 2014 cited in Ofulue & Benyoucef, 2022) and data wrapping (Hunke et al., 2022; Ofulue & Benyoucef, 2022). We, however, exclude these models from our investigation as we focus on direct spendings of governments.

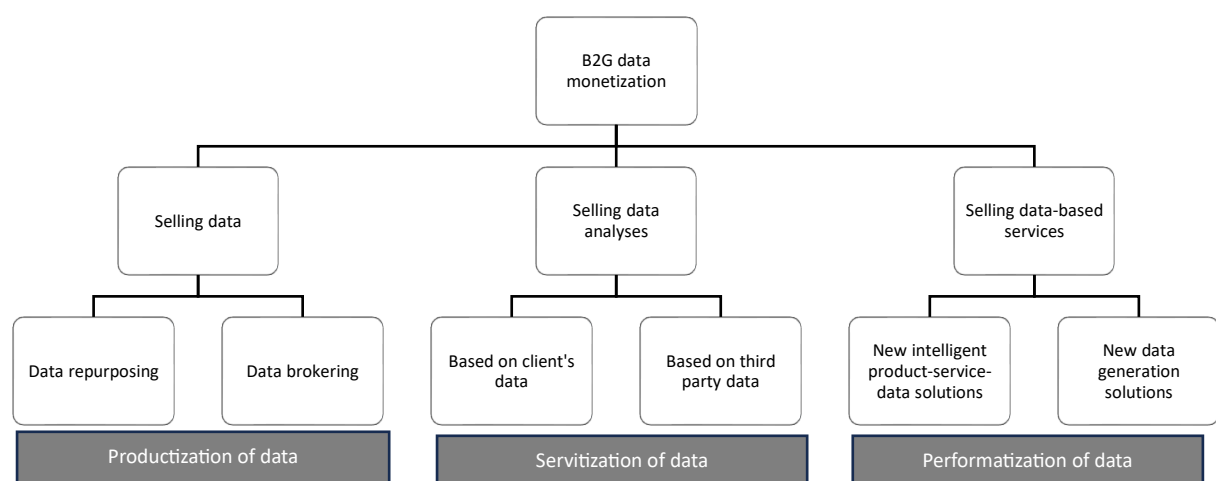


Fig. 1. Prospective routes to Business-to-Government (direct) data monetization

Taking this categorization as a starting point, our study is the first systematic effort to collect empirical evidence on the practice of data purchasing by governments and data monetization by companies in the Netherlands. Academic research on B2G data sharing has so far overlooked the data purchasing model and almost solely focused on voluntary access to private sector data.

3. Methodology

We conducted a deep dive into the (open) government data relevant to government spending to collect publicly available information on our research question. To do so, we collected data on data purchasing ‘cases’ and focused on mapping which governments purchase private sector data from whom, what type of data is purchased, for what purpose, and how. We used publicly available information provided by two platforms to assemble our dataset: TenderNed¹, and OpenRaadsinformatie². Two other platforms, OpenSpending and Findo, were consulted but did not deliver results at the required level of granularity. The data collection took place between August 2023 and December 2023.

The first platform TenderNed serves as the Dutch government's official procurement platform, facilitating the dissemination of tender notifications by public authorities. Tenders exceeding certain monetary thresholds are subject to European procurement procedures as outlined in the Tender Act of 2012. For central authorities, such as ministries and judicial bodies, the threshold for service supply is €143,000, while for decentral authorities like municipalities, provinces, and water boards, it is €221,000. Below the European threshold, governments may follow a national procedure. The publicly available database, TenderNed 2016-2023 (Q1-Q2), provides all European and national tenders published between 2016 and 2023. We analyzed the dataset using several Common Procurement Vocabulary (CPV) codes, these codes are assigned to all sorts of public contracts for services, supplies, and works. We collaborated with a data analyst from TenderNed to select the following codes for conducting our query: data processing (72300000-8), database and operating software (48600000-4), data-base services (72320000-4), data processing services (72310000-1), data collection services (72313000-2), and data transmission services (72318000-7). The search yielded 259 tenders. Subsequently, an in-depth and manual analysis of all provided documentation of the tenders was necessary to ascertain whether it was in the scope of our research. This resulted in 14 tenders on a national, regional, and local level.

The second platform, OpenRaadsinformatie, is an initiative provided by the NGO Open State Foundation in collaboration with the Association of Dutch Municipalities (VNG). The platform provides information by extracting governmental board meetings and other documents from information systems of decentral authorities. The platform functionality did not allow us to conduct a systematic search to identify cases of data purchasing. Therefore, we used a workaround technique and conducted a search based on a list of 95 potential providers of data or data services. This list of providers was manually compiled based on in-sight during our talks with experts in the field, our previous research, a search on participants of various GovTech conferences, the TenderNed dataset, among others. This list is made available as open research data and can be accessed here³. Like the search within TenderNed, an in-depth and manual analysis of all provided documentation was

¹ <https://www.tenderned.nl/>

² <https://zoek.openraadsinformatie.nl/>

³ <https://doi.org/10.24416/UU01-YATO7Q>

necessary to ascertain whether the cases fall within the scope of the research. The search eventually resulted in 15 additional use cases.

Finally, we have collected several additional use cases through an open call advertised in a number of relevant professional networks, such as the knowledge network of VNG on data and society (Kennisnetwerk Data & Samenleving), and through talks with 8 experts in our network (e.g. municipal privacy officers). This resulted in 4 additional cases.

In total, we ended up with **33** cases. The complete overview of the cases has been made available as open research data and can be accessed here⁴. In terms of data analysis, for each case, we identified the buyer and seller, the type of data/offering, the procedure, the domain and provided a short description of the case.

4. Results

4.1. Who: buyers and sellers

Based on our analysis, we identified local government authorities, regional authorities, as well as national level agencies, as buyers of data and data services in our sample. Figure 2 shows a visualization which illustrates the connections between the private sector providers of data and services in the Netherlands and the buyers – the governmental entities based on the collected cases. To the left of the figure, the governmental entities are delineated: (A) national governments, (B) regional governments, and (C) local governments. In certain instances, two distinct cases have been identified per governmental entity, resulting in their duplication. The connectivity between governments and providers is visually depicted by lines, denoting the procurement of data/services from certain providers. Each governmental entity represents a single case; where multiple providers are connected to a single entity, it signifies the extent of provider involvement within a case. The colors assigned to entities denote the policy domains to which the data/services belong within governments. The domains are explained in detail in section 4.2. Figure 2 excludes three cases connected to National Road Data Portal (NDW) to maintain readability of the figure. NDW is a collaborative initiative with its dedicated implementation body, facilitating governmental cooperation in the acquisition, dissemination, and utilization of data aimed at addressing challenges in mobility and public infrastructure. Each case connects to numerous data/service companies (29 providers identified).

The landscape of private sector providers of data and data services in the Netherlands appears to be rather versatile, including data science/analytics companies (e.g. Dat.mobility, Resono, Mezuro), providers of data, software, consultancy services (e.g. Atos, Bureau de Groot Volker, ECA International), and foundations (e.g. Stichting LISA, Pensioenfonds Zorg en Welzijn).

A number of companies appeared more frequently in our sample (e.g. Cyclomedia, Dat.mobility, Connection Systems, Vialis) which signals that there is an established relation between a number of governments and certain companies (e.g. Municipality of Vijfheerenlanden, Province of North Brabant, Province of Utrecht, NDW) and that these providers seem to have established a solid position on the data market. For instance, Dat.mobility has provided data and/or related services to multiple governments (e.g. the Municipality of Lindewoord, Province of North Brabant and Limburg). Interestingly, the company collaborates often with other providers within a certain project, including data providers like Mezuro and Mobidot. Another example is Connection Systems which is specialized

⁴ <https://doi.org/10.24416/UU01-YATO7Q>

in the collection, processing, and presentation of data and images for various applications in security, surveillance, and law enforcement. One of their frequently encountered services involves generating traffic-related data through license plate registration, utilizing sources such as Automatic Number Plate Recognition (ANPR) cameras. The company offers comprehensive solutions, including the provision of cameras, installation, monitoring, and data analytics. Within our sample, Connection Systems primarily serves national and regional governments, such as the NDW.

Another interesting finding in our sample is the presence of the non-profit organizations. For instance, the foundation Stichting LISA provides a database containing information on all establishments in the Netherlands where paid work is conducted. They offer access to both open data and data for a fee. Our sample reveals that North Brabant purchased data from Stichting LISA to inform the development and improvement of economic policies. Another non-profit company that appeared in our sample is Pensioenfonds Zorg en Welzijn (PFZW), one of the largest pension funds in the Netherlands. Our sample reveals that PFZW delivers labor market data to the Ministry of Health, Welfare, and Sport (VWS) to address labor market challenges in the healthcare and welfare sector.

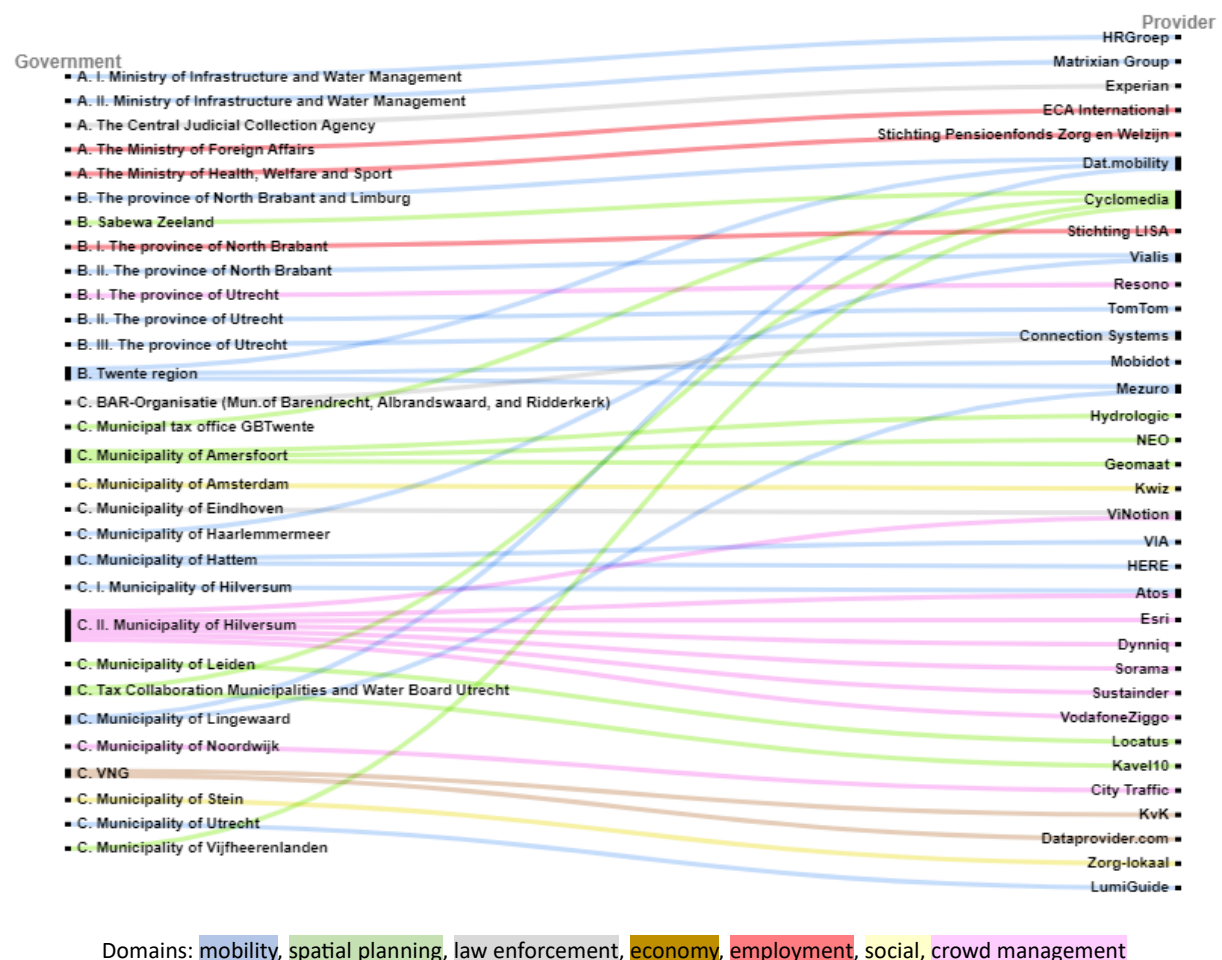


Fig. 2. Company providers and government buyers of data, insights, and data-driven services in the Netherlands

4.2. What and how: types of data monetization

When analyzing the collected cases, we encountered that categorizing some of them based on the three data monetization routes proved to be challenging. This is due to the insufficient level of detail about the cases that we could derive based on publicly available documentation. Cases that involved provision of a dashboard or monitor were categorized as selling data analytics, or servitization. In Table 1, the cases are presented in the format “Company provider: data/service provided (government organization)”.

In our study, we found evidence that in B2G context, productization of data takes place in various domains (e.g. mobility, economic, law enforcement) and at all levels of government (local, regional, national). Among the cases we identified (7 cases), we find various players selling data: data brokers (Experian), foundations (e.g. Stichting LISA), data aggregators (e.g. Dataprovider.com). We did not find many cases of data re-purposing towards government organizations, but this might have to do with limitations of our search strategy. Other providers, like Stichting LISA, offer some datasets for free and some for a fee.

In our research, we also identified a number of cases (12 cases) illustrating the servitization of data (selling analyses, instead of data). This category of cases demonstrates that, alongside data, governments, at different levels, are procuring data analytics solutions to, for instance, address such societal issues like crowd management, managing tourism, population growth and spatial development. For instance, ECA International provides data and analysis on the cost of living across diverse global locations and assists organizations in devising equitable compensation packages for their employees stationed abroad, which they provided to the Dutch Ministry of the Interior and Kingdom Relations (BZK). In many cases, the analyses are based on data combined from various sources (e.g. offerings by companies like Locatus). For instance, the company Mezuro, specialized in geospatial data, provides data collections based on Internet of Things (IoT) sensors, satellite imagery, GPS data from mobile devices, open data sources, and crowdsourced data (e.g. location-based information). Mezuro data is often used by data analytics companies, such as Dat.mobility.

A number of cases (7 cases) we collected fall under the model of selling data-driven services. Here we observe that all identified cases belong to the mobility domain. This could point to the existence of an active data ecosystem but could also result from our search strategy (partly based on a predefined list of providers). In this category, we find a number of cases where a private provider offers a product-service-data solution and, in some cases, helps the government to act on insights. For example, the province of North Brabant contracted the company Vialis to provide traffic installations together with data analytics solution analyzing traffic signals. Another example is the Schwung-app, marketed by Vialis and Infoplaza (known for Weerplaza), which collects data to optimize traffic signal timing for cyclists based on GPS data while using the app. Subsequently, this data is transmitted to Vialis for potential further mobility analysis. The practice of gathering traffic data, such as location data, via mobile applications, is increasingly common.

We also categorized a number of cases as possible combinations of various routes (see Table 1). Some providers in this category offer a range of services. For instance, Cyclomedia collects visual data in public spaces for many governments using patented vehicle-mounted camera systems, producing 360° panoramic images and LiDAR data, alongside additional services like oblique aerial photography. Their services are in demand among numerous governments, who often opt for joint procurement to leverage cost efficiencies (e.g. BGTwente). However, in recent years, competitors like Kavel10 have also entered the market, this company is also represented in our sample.

We also found that data or data services are purchased via different procedures, such as tendering (mostly at the national and regional level) or private contracting (mostly at the local level, for amounts lower than 221,000 euro). For instance, municipalities often establish long-term contracts with data services companies, allowing them to access various services for pertinent projects within the municipality. As an illustration, our sample indicates that the Municipality of Haarlemmermeer utilized the Schwung app and its associated services (e.g. the conversion of nineteen traffic signal installations) under an already existing contract with Vialis.

Table 1. B2G data monetization routes based on the collected cases

Domains: mobility, spatial planning, law enforcement, economy, employment, social, crowd management

* partially government funded

| Selling data | Selling data analyses | Selling data-driven services |
|---|--|--|
| [4] Stichting LISA: employment data (Pr North Brabant) | [2] HRGroep: traffic signs data analytics (IenW) | [1] Atos: Smart City platform (M Hilversum) |
| [6] PFZW: labor market data (VWS) | [5] ECA International: cost of living analytics (BZK) | [7] Vialis: traffic signals analysis service (Pr North Brabant) |
| [30] DataProvider.com/KvK: call detail records (4 municipalities) | [8] Kwiz: customer data analytics (M Amsterdam) | [10] Connection Systems: traffic behavior service (Pr Utrecht) |
| [33] Experian: phone numbers data (CJIB) | [12] Dat.mobility: mobility dashboard and analytics (Pr North Brabant) | [11] Connection Systems: traffic camera service (3 municipalities) |
| [3] Multiple providers: floating car data (IenW NDW) | [20] Mezero/Dat.mobility: mobility analyses (M Lingewaard) | [15] Vialis-Infoplaza: traffic light optimization (M Haarlemmermeer) |
| [9] Multiple providers: sensor/telraam data (IenW NDW) | [21] Locatus: retail analytics (M Leiden) | [16] LumiGuide: bicycle parking optimization (M Utrecht) |
| [14] Multiple providers: cycling data (IenW NDW) | [28] VIA/HERE: traffic safety monitor (M Hattem and North Veluwe region) | [22*] Multiple providers: mobility-as-a-service (Twente region) (Netmobil project) |
| | [23] City Traffic: crowd density dashboard (M Noordwijk) | |
| | [25] TomTom: floating car data analytics (Pr Utrecht) | |
| | [13] Matrixian Group: logistics/trucks data analytics (IenW) | |
| | [26] Atos/Esri: crowd monitor (M Hilversum) | |
| | [27] Resono: crowd monitor (Pr Utrecht) | |
| | [29] Zorg-lokaal: dashboard social domain (M Stein) | |
| [17] Cyclomedia: aerial photos (M Twente) | | |
| [19] Cyclomedia: street lidar (M Vijfheerenlanden) | | |
| [24] Cyclomedia: cycloramas (Sabewa Zeeland) | | |
| [18] Kavel 10/Cyclomedia: aerial photos (9 municipalities and 1 waterboard) | | |
| [31*] Atos/Intel: safety incident prediction dashboard (M Eindhoven) (City Pulse project) | | |
| [32*] Multiple providers: spatial insights for neighborhood (M Amersfoort, RIVM) (LIVE project) | | |

PFZW = Pension fund, VWS = Ministry of Health, Welfare and Sport, KvK = Chamber of Commerce, CJIB = Central Judicial Collection Agency, IenW = Ministry of Infrastructure and Water Management, BZK = Ministry of the Interior and Kingdom Relations

4.3. What for: policy domains

We established that data or data services are purchased in the following policy domains: economic affairs (which encompasses matters spanning the economy, culture, and the labor market), crowd management (focusing on specific projects aimed at regulating visitor flows within public areas, applicable across various domains), the social domain (in charge of youth care, long-term care, and income support), employment (addressing labor market issues, sometimes overlapping with economic or social affairs), spatial (planning) domain (responsible for maintaining, upholding, and managing the physical environment for citizens and businesses), law enforcement (ensuring public order and safety), and mobility do-main (handling road management, traffic regulation, parking, and transportation). These domains were identified deductively, drawing on Vermeulen's (2015) categorization of social policy decentralization in the Netherlands, supplemented by our collective expertise across various domains and departments within both central and decentral Dutch governmental entities.

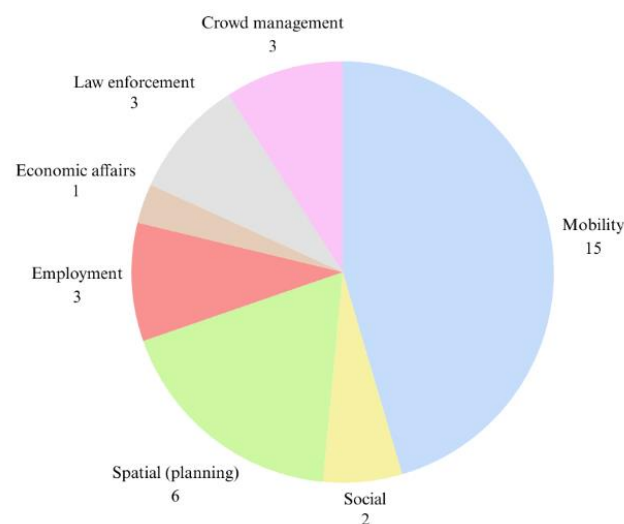


Fig. 3. Policy domains in which purchasing of data, insights, and data-driven services occurs in the Netherlands

Figure 3 provides an overview of the various domains for which we could identify purchases of data, data insights or data-driven services. The domains where most of our identified use cases are concentrated are mobility (15 cases), spatial (planning) domain (6 cases), law enforcement (3 cases), crowd management (3 cases), and employment (3 cases).

Mobility is one of the domains where there is an active data market. In general, mobility challenges arise at local, regional, and national levels. However, it is common that mobility challenges are addressed by regional and national governments, which is also reflected in our sample (e.g. IenW, NDW, province of North Brabant). Based on our sample, we found various types of data/services on offer: data-driven services, such as smart city platforms, data insights, such as traffic analyses or calculation models, as well as datasets, such as Probe Vehicle Data, Floating Car Data, and crowd-sourced data are prevalent. Providers that specifically offer many services in this domain are the

aforementioned Dat.mobility and Vialis; they provide both data-driven services and data insights as well as datasets, and something even in combination.

The *spatial (planning)* domain represents another domain where an active data market exists. This domain involves tasks, such as the oversight of public spaces, including the maintenance, management, and regulation of amenities, as well as the control of visitor flows and the upkeep of cleanliness. These responsibilities are usually undertaken at the local level, which is also reflected in our sample. Based on our sample, we found various types of data/services on offer, however, the most common is the procurement of data analysis and insights, as well as combination of data monetization routes. A prevalent provider catering to various local governments in this domain is the aforementioned Cyclomedia. For purposes such as spatial planning, Cyclomedia offers, among other services, Street View imagery and oblique aerial photographs.

Within the *law enforcement* domain, several cases have been identified. The oversight of public order and safety constitutes a shared responsibility between central and decentralized authorities. Within municipalities, this jurisdiction is specifically allocated to the domain of public order and safety. At the national level, this entails the involvement of implementing agencies, such as the Central Judicial Collection Agency (CJIB) and law enforcement agencies (e.g. the police, and customs), operating under the Ministry of Justice and Security (J&V). Both governmental levels are reflected in our sample. Our analysis reveals the presence of both data-driven services and the procurement of datasets within our sample, as well as a combination of data monetization routes. An illustrative ex-ample of a prevalent service within this domain, drawn from our sample, includes the procurement of ANPR cameras for the enforcement of traffic regulations. A predominant provider of this service is the firm Connection Systems.

Crowd management is in fact not formally recognized within governmental structures or domains. Nonetheless, for this study, we have grouped these cases due to their significant overlap and prominence. Typically, such cases would fall within various established domains, depending on the objectives of a project. Crowd management entails the analysis of visitor flows within a specific area, such as festival grounds, sports stadiums, or commercial districts. This under-scores why crowd management activities primarily occur at the local level. The objectives of crowd management initiatives can vary widely, ranging from ensuring public safety and order to stimulating economic activity or enhancing market-ing strategies. Within our sample, the predominant offerings are data analyses, for instance, the Crowd Monitor implemented by the Municipality of Hilversum and provided by Atos, generating data through sensors and GPS data to map visitor flows. Another example is the well-known Dutch company Resono, which provides location data generated through various (Dutch) apps, such as weather apps, to map visitor numbers and crowd density, among other services.

Finally, within the domain of *employment*, several cases have been identified. The oversight of employment is a collaborative effort between central and decentralized authorities. At the municipal level, employment is predominantly categorized under the social domain. On a national level, these responsibilities are as-signed to ministries, such as the Ministry of Health, Welfare, and Sport. Within our sample, predominantly regional and national governmental bodies have been represented in this domain. Our sample underscores the significance of procurement of datasets and data insights within this domain.

5 Discussion

Previous studies in the EU have noted that governments procure data (Micheli, 2022), although this phenomenon lacks systematic investigation. Concurrently, the data market, especially in the B2G sector, is rapidly evolving and presents distinct challenges, as outlined in section 1 (Martens & Duch-Brown, 2020). Our research aims to provide deeper insights into the who, what, how, and what for governments procure data, as detailed in section 4.

A first important overarching finding is that there is a **lack of transparency** about which data is bought by whom, how and what for. For instance, there is no centralized database tracking which entities purchase what data, hindering a comprehensive understanding of the scope and expenditure of data procurement. A large portion of such transactions of interest are scattered at the level of project invoices and thus difficult to uncover in a systematic way. As far as we know, this problematic lack of transparency is also present in many other countries.

The lack of transparency created some challenges for our research. Nonetheless, we contribute several insights to the literature on government data procurement practices. Our findings reveal that, looking at the example of The Netherlands, governments at all levels (national, regional, local) engage in various approaches to data acquisition that align with three data monetization routes of companies: productization of data, servitization of data, performatization of data, and combinations thereof. Although our data does not allow for any conclusions regarding why certain governments choose to buy data, data analyses, or data-driven services (“sourcing decision”), we uncovered that in our sample data-based services appear largely in the mobility domain. At the same time, we did not find any cases where a big tech company would be a supplier which once again signals the limits of relying just on publicly available evidence. Moreover, our analysis uncovers several issues. Chief among these are market complexities and ethical considerations. Regarding **market complexities**, a diverse landscape of data providers offers a wide array of data types. Some providers specialize in niche offerings, complicating competitive assessments for procurers. Conversely, certain providers wield considerable market dominance, posing challenges for fostering fair competition as well. Furthermore, data procurement entails specialized models that might not be universally understood by all procurers, particularly within smaller governmental bodies, raising doubts about the consistency of price-quality ratios.

There are also **ethical considerations** to address. Some data originate from questionable practices, such as data collected from a service to warn automobilists of speeding cameras. Should governments purchase data that were collected with the intention to support circumventing the law? Other services might thrive on originally public data and generating profits by selling them back to public organizations.

To mitigate these challenges, enhancing transparency about government data (service) procurement practices and fostering joint, more professional procurement approaches are proposed solutions. The significance of these themes in public procurement has been underscored in broader literature (e.g. Bauhr et al., 2020; Schotanus, 2022), and also seem particularly relevant to procurement of data, data insights, and data services. It would be beneficial for taxpayers and public organizations to critically review practices of such purchasing. Economic arguments aside, there are also political and ethical arguments to be made. The practice of data purchasing risks putting local smaller governments, which typically have limited IT and procurement capacity, in a disadvantageous position vis-à-vis the private sector. With an eye to public spending and sovereignty of public institutions such as municipalities, provinces or the national government, more insights into procurement practices and an overview of the marketplace could support collective bargaining (Carrera et al., 2021), and the curtailing of

monopoly positions. For instance, in our sample of cases, we found limited collaboration among Dutch municipalities in purchasing data, data analyses or data-based services. Ideally, this would extend to a European level where cities, regions, or member states could work together in sharing information, advertising their tenders, procuring jointly data and related services, and making these data available on national levels.

Another beneficial outcome of more transparency could be an evaluation of data-driven products that have been purchased by many government organizations. It might lead to the conclusion that many dashboards are hardly used and could be falling short of providing actionable information. Such an evaluation and comparative overview of costs and benefits of data, insights from data and data-driven services could inform expected quality, applicable use cases, and procurement and governance. Additionally, it might stimulate innovation and competition in the marketplace.

6 Conclusion

In this exploratory study, our objective was to map the state of play as regards purchasing of data, data insights, and data-driven services by governments from the private sector for addressing societal issues (who, what, how, and what for). To answer this question, we investigated the example of The Netherlands and collected cases (n=33) of such purchases at national, regional, and local levels of government. We analyzed these cases based on the parties involved, the content of the purchase, the purpose, the procedure used. We also proposed a conceptual framework based on the literature of business-to-government data monetization routes which, as demonstrated in our study, can be useful as an analysis tool for uncovering governments' purchases of data from private companies. Our study was the first systematic attempt to map data purchasing at the level of a country and we encourage other researchers to replicate this effort in other countries, be it in or outside the EU.

Our main conclusions are that data are purchased in different forms. There is however no comprehensive overview of the procurement of data, insights from data and data-driven services available, which makes this a rather opaque area of public spending. Such knowledge could be relevant for reducing information asymmetry in this particular marketplace, controlling public spending, and considering possible vendor dependencies.

The limitations of our study have to do with the search strategy which we were forced to employ under the conditions of low transparency and absence of granular data. Thus, our results should be interpreted as an exploratory overview, rather than an exhaustive mapping of purchases by Dutch governments. Future research can investigate the collected cases in more depth and broaden the data collection to identify more cases. Such issues as the business models of companies in the B2G market could be an interesting research angle. Studying how other EU member states procure data and assessing the impacts of any notable differences in market approaches would also be insightful. Future research could also dive deeper in the processes and challenges experienced by government organizations when it comes to purchasing data, data insights or data-driven services and how these practices can be further professionalized.

Data statement

The data underlying this research (the dataset of cases and the list of providers) has been made openly available and can be accessed here: <https://doi.org/10.24416/UU01-YATO7Q>

Funding

This research has been funded by the Dutch Research Council (NWO) under the Open Competition XS scheme grant number: 406.XS.03.053

References

- Bauhr, M., Czibik, Á., de Fine Licht, J., & Fazekas, M. (2020). Lights on the shadows of public procurement: Transparency as an antidote to corruption. *Governance*, 33(3), 495-523.
- Borowitz, M. (2019). Government data, commercial cloud: Will public access suffer?. *Science*, 363(6427), 588-589.
- Carrera, P., Katik, S., & Schotanus, F. (2021). Joint procurement of complex products: Actual price savings, perceived nonmonetary advantages, disadvantages and impediments. *Journal of Public Procurement*, 21(2), 167-182.
- Christl, W., & Spiekermann, S. (2016). Networks of control. A report on corporate surveillance, digital tracking, big data & privacy Facultas.
- George, J. J., Yan, J., & Leidner, D. E. (2020). Data philanthropy: corporate responsibility with strategic value?. *Information Systems Management*, 37(3), 186-197.
- Günther, W. A., Mehrizi, M. H. R., Huysman, M., Deken, F., & Feldberg, F. (2022). Resourcing with data: Unpacking the process of creating data-driven value propositions. *The Journal of Strategic Information Systems*, 31(4), 101744.
- Helmond, A., & Van Der Vlist, F. (2023). Situating the Marketization of Data. In Amsterdam University Press eBooks (279–286). <https://doi.org/10.1515/9789048555444-020>
- Hunke, F., Heinz, D., & Satzger, G. (2022). Creating customer value from data: foundations and archetypes of analytics-based services. *Electronic markets*, 1-19.
- Jarvenpaa, S. L., & Markus, M. L. (2020). Data sourcing and data partnerships: Opportunities for IS sourcing research. *Information Systems Outsourcing: The Era of Digital Transformation*, 61-79.
- Kotlarski, J., Oshri, I., Dibbern, J., & Mani, D. (2018). MIS quarterly research curations. *IS Sourcing*
- Krasikov, P., Eurich, M., & Legner, C. (2022). Unleashing the Potential of External Data: A DSR-based Approach to Data Sourcing. In ECIS.
- Langseth, M., Data-Driven Decision-Making in the Public Procurement System: The Case of Norway, PhD dissertation, NTNU, 2024.
- Martens, B., & Duch-Brown, N. (2020). The Economics of Business-to-Government Data Sharing (February 19, 2020). JRC Technical Report, JRC Digital Economy Working Paper 2020-04, Available at SSRN: <https://ssrn.com/abstract=3540122>

This is a working paper submitted and under peer review in an academic journal. Previous version of this research has been presented at Overheid 360 Congress on 19 June 2024 in Utrecht, The Netherlands.

Mayer-Schönberger, V., & Ramge, T. (2022). *Access Rules. Freeing data from big tech for a better future.* University of California Press.

Micheli, M. (2022). Public bodies' access to private sector data: The perspectives of twelve European local administrations. *First Monday*.

Ofulue, J., & Benyoucef, M. (2022). Data monetization: insights from a technology-enabled literature review and research agenda. *Management Review Quarterly*, 1-45.

Parvinen, P., Laitila, M., Pöyry, E., Gustafsson, R. & Rossi, M. (2020). 'Advancing data monetization and the creation of data-based business models', *Communications of the Association for Information Systems*, vol. 47, no. 2, 38, pp. 25-49. <https://doi.org/10.17705/1cais.04702>

Ritala, P., Keränen, J., Fishburn, J., & Ruokonen, M. (2024). Selling and monetizing data in B2B markets: Four data-driven value propositions. *Technovation*, 130, 102935.

Ruppert, E., Isin, E., & Bigo, D. (2017). Data politics. *Big data & society*, 4(2), 2053951717717749.

Schneider, C. O., Bremen, P., Schönsleben, P. and Alard, R. (2013) "Transaction cost economics in global sourcing: Assessing regional differences and implications for performance", *International Journal of Production Economics*, 141(1), pp. 243–254.

Schroeder, R. (2016). Big data business models: Challenges and opportunities. *Cogent Social Sciences*, 2(1), 1166924.

Schotanus, F. (2022). Open public procurement data by default, recommendations for the Dutch central government for making different types of public procurement data publicly available. Center for Public Procurement, Universiteit Utrecht, Utrecht.

Sharon, T. (2020). Beyond hostile worlds: The multiple sphere ontology of the digitalization and Googlization of health. Keynote lecture. Conference on Digital Health: Ethical, Legal and Social Perspective for a Sustainable Future, 9 December 2020.

Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world.* Oxford University Press.

Verhulst, S., & Sangokoya, D. (2015). Data collaboratives: Exchanging data to improve people's lives [Medium]. The GovLab.

Vermeulen, W. (2015). Decentralisation of social policy in the Netherlands. *Decentralisation of education, health and social protection: Issues and challenges*, 127.

VNG Realisatie (2017). *Koken met data.* Publication of the Association of the Dutch Municipalities (VNG). Accessed on 23-11-2023 from <https://publicaties.vngrealisatie.nl/2017/koken-met-data/de-ingredienten/>