

Measuring public innovation in the EU: the STARPIN methodology

2019 edition



**Measuring public
innovation in the EU:
the STARPIN methodology**

2019 edition

Manuscript completed in March 2019

Printed by the Publications Office in Luxembourg

The Commission is not liable for any consequence stemming from the reuse of this publication.

Luxembourg: Publications Office of the European Union, 2019

© European Union, 2019

Reuse is authorized provided the source is acknowledged.

The reuse policy of European Commission documents is regulated by Decision 2011/833/EU (OJ L 330, 14.12.2011, p. 39).

Copyright for the photographs: Cover © Dostoevsky /Shutterstock

For any use or reproduction of photos or other material that is not under the copyright of the European Union, permission must be sought directly from the copyright holders.

For more information, please consult: <https://ec.europa.eu/eurostat/about/policies/copyright>

| | | | | |
|--------|------------------------|----------------|---------------------|-------------------|
| PDF: | ISBN 978-92-76-01095-1 | ISSN 2315-0815 | doi: 10.2785/215876 | KS-GQ-19-002-EN-C |
| Print: | ISBN 978-92-76-01097-5 | ISSN 2363-197X | doi: 10.2785/708816 | KS-GQ-19-002-EN-N |

Contents

| | |
|--|-----------|
| Contents | 3 |
| Abstract | 5 |
| 1. Introduction | 6 |
| 2. The StarPIN approach | 8 |
| The conceptual pillars of the StarPIN approach | 9 |
| Placing public service innovation at centre stage..... | 9 |
| Adopting an object-based approach..... | 9 |
| Interpreting innovation as an increase in value creation | 11 |
| A new hierarchy of data sources | 13 |
| 3. The foundations of the StarPIN approach | 16 |
| Combining insights from complementary strands of literature | 16 |
| Approaches to public service innovation | 17 |
| Key streams of literature on public service innovation | 21 |
| The literature on public service dominant logic | 21 |
| The literature on ‘value creation and co-creation’ | 22 |
| 4. Towards an analytical framework: topics, issues and variables | 26 |
| 5. A classification grid | 28 |
| The identification of the target statistical unit: public institutions | 28 |
| Which are the conditions for adopting an ‘object’ approach to measure public innovation? | 32 |
| The services of general interest as statistical units: a CPA-based classification | 36 |
| 6. Methodology for data collection and data analysis | 39 |
| The process | 39 |
| Current status of site-centric measurements | 43 |
| Application of the site-centric measurement in two countries | 45 |
| Pilot service waste collection..... | 45 |
| Keyword identification strategy..... | 48 |
| Web-scraping process..... | 49 |
| Validation strategy | 49 |
| Summarising and integrating collected data..... | 50 |
| Table 9: A simplified template structure to integrate data on public sector innovation in e.g. the waste management domain | 55 |
| Pilot service Health homecare..... | 56 |
| 7. Main results of the methodology application | 62 |
| Pilot service waste collection in the five pilot countries | 62 |
| Scaling up the indicator generated | 67 |
| Cross-sectional analysis of data | 76 |

| | |
|---|------------|
| Preliminary results for the pilot service Health homecare | 76 |
| 8. Learning from the application of the methodology | 80 |
| Main difficulties..... | 80 |
| Lessons learnt through the testing activity | 81 |
| Directions for future improvements..... | 82 |
| References | 87 |
| Annex 1: Classification Grid | 96 |
| Annex 2: User manual | 112 |
| Annex 3: Keywords and keyword baskets..... | 116 |
| Annex 4: Example of data for validation and cross-sectional analyses | 130 |
| Annex 5: Calculation of regional and national averages of innovativeness levels..... | 144 |

Abstract

This work accounts for, and synthesises the main results of, the **StarPIN project — Statistical Reporting on Public Innovation** — developed by the **Dipartimento di Economia, Società, Politica (DESP)** of the **University of Urbino, Italy**, with the support of **dialogic** — innovatie. interactie, Utrecht, the Netherlands. The StarPIN project was supported by Eurostat's Unit working on innovation statistics. The StarPIN team was composed of Annaflavia Bianchi, Claudio Cozza, Giovanni Marin, Robbin Te Velde, Antonello Zanfei, and Emy Zecca. Gregor Kyi and Giulio Perani were the Eurostat project managers and Stefania Panaitescu ensured editing and support.

The project proposes a theoretical and a methodological framework for enriching the measurement ability of public sector innovation, and discusses the implications for data collection and analysis. A preliminary test of the framework is carried out by means of pilot applications to specific public services.

The project focuses on public service innovation as a key locus of value creation for society, choosing an object-based rather than the more commonly adopted subject-based approach. Issues concerning the measurement of public service innovation are addressed in statistical terms and consistently with the official public sector functions classifications. The approach places particular emphasis on web-scraping to capture the technological level of websites used by public administrations, and to evaluate the degree of innovativeness of specific services within selected public functions. Data collected through web-scraping can be combined with administrative data at the level of individual services. The paper presents pilot applications to specific public services in a limited number of countries.

The paper is structured according to the following steps: First a theoretical framework is sketched for the analysis of innovation in the public sector, focusing on service innovation and public value creation. Second, the consistency between public functions, public institutions and public service classifications is discussed. Third, the indicators of innovativeness are generated for selected public services. Fourth, the paper illustrates the main lessons learnt from the pilot application of the methodological framework, and draws implications for future research. A user manual is annexed to the paper for replicating the pilot tests described in the paper.

Editors

Annaflavia Bianchi, Claudio Cozza, Giovanni Marin, Giulio Perani, Robbin Te Velde, Antonello Zanfei, Emy Zecca

Eurostat Project Manager

Gregor Kyi

Production

Stefania Panaitescu

Citation

Bianchi A., Cozza C., Marin G., Perani G., Te Velde R., Zanfei A., Zecca E. (2019) Measuring public innovation in the EU: the STARPIN methodology. 2019 Edition, Eurostat, Manual and Guidelines, Luxembourg: Publications Office of the European Union, ISBN 978-92-76-01095-1 ISSN 2315-0815 doi: 10.2785/215876 KS-GQ-19-002-EN-C, available at: <https://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/KS-GQ-19-002>

Keywords: *public sector innovation, public services, public value creation, web-scraping, administrative data, measuring innovation, data-source integration, object-based approach*

1

Introduction

Innovation is central for the progress of economies and societies as it helps to sustain and increase their competitiveness and welfare. There are established guidelines and tools to measure innovation in businesses, in particular the 'Oslo Manual – Guidelines for collecting, reporting and using data on innovation' and the 'Community Innovation Survey', which increases the harmonisation of surveys on business innovation in the European Union since 1992. However, also innovation in the public sector is increasingly in the focus. There is an increasing demand for benchmarking the efficiency and quality of public services as well as for identifying the factors that contribute to desirable outputs and outcomes of innovation in the public sector.

The 4th edition of the Oslo Manual acknowledges the need for a general definition of 'innovation' that is applicable in all institutional sectors or entities (including the public sector) and which retains consistency with the definition of 'business innovation'. It provides this definition: "An innovation is a new or improved product or process (or combination thereof) that differs significantly from the unit's previous products or processes and that has been made available to potential users (product) or brought into use by the unit (process)." The Manual refers to specific features of innovation in the general government sector, and points to the recently growing demand for studies within the government and the public sector in general. Such studies have adapted the guidelines in the previous editions of the Manual to develop surveys on innovation in public sector administration organisations, or have added questions that are explicitly designed for the public sector. In addition, the fourth edition of the Oslo Manual extends its guidance on the usage of complementary data sources and collection methods, including administrative sources and big data, and points to an integrated use of sources and methods to address user needs.

Though there are surveys on public sector innovation in some EU Member States that do indeed merit attention, the European Statistical System (ESS) does not yet have any sustained and shared methodology or tool to report on innovation in the public sector.

This publication is driven by the need within the ESS to work more on the statistical measurement of public sector innovation. However, it is very far from providing *the* one and only solution. In contrast, the paper shall be read as a collection of principal considerations, ideas and first experiences that may trigger a debate and, in the long run, may help to advance towards common methodology in the area. It does not see itself in competition to, to replace or amend existing surveys on public sector innovation, existing statistics on the subjects of the presented pilot studies, nor existing guidelines in the area. However, it may broaden our thinking as it reaches out from the established context, methodology and classifications of official statistics, towards innovative methods and tools of potential future data collections. This publication can help analysts, subject matter experts, policy-makers and statisticians to reflect on possible demands and usages of statistics about the heterogeneous area of public sector innovation, as well as about modern, efficient and 'fit-for-purpose' tools and technologies that might meet such demands in the future.

In that view, Eurostat has launched and managed an 18-month Project (from January 2017 to June 2018) on Measuring Public Innovation in the EU, also known by the acronym StarPIN. A team of professors and doctoral students from the Università degli Studi di Urbino Carlo Bo, Dipartimento di Economia, Società, Politica (DESP), took the discussion on measuring Public Sector Innovation (PSI) and designed a novative approach and methodology to measure PSI.

Steered by Eurostat's Science, Technology and Innovation statistics team, the StarPIN scholars and experts have set a framework for measuring the phenomenon.

The paper is organised as follows. Chapter 2 sketches the conceptual pillars of the StarPIN approach that is: placing public service innovation at centre stage, adopting an object-based approach and identifying public innovation as an increase in public value. It draws implications for statistical analysis and proposes a new hierarchy of statistical sources to be used to measure public innovation. Chapter 3 discusses the theoretical foundations of the StarPIN approach, by illustrating and combining two main streams of literature — the service-dominant logic and theories of public value creation. Chapter 4 briefly sketches the topics and issues at stake when measuring public innovation and identifies the key variables that need to be measured. Chapter 5 proposes a classification grid that enables the linking of service activity codes (derived from CPA classification) with COFOG government functions and with NACE codes to identify the set of subjects involved in innovative activities. Chapter 6 illustrates the process through which services are identified and characterised in terms of innovativeness, and through which data are collected and organised. Using two public service domains as illustrations of this data search method, the following steps are described: the identification of public services, in correspondence with extant classifications, and of institutions involved in their provision; the description of service characteristics that can be associated with different levels of innovativeness; the singling out of descriptive tools (keywords and 'baskets' of keywords) that can be handled via web-scraping techniques; the running of web-scraping using appropriate techniques; the organisation of results of data obtained through the web-scraping process; the validation process through the use of complementary data sources, including administrative data. Chapter 7 discusses the main results stemming from the pilot application of the StarPIN framework and methodology to five EU countries, with illustrative analyses of newly collected data. Chapter 8 reflects on difficulties encountered in the data collection and data analysis, identifies the main lessons learnt through the pilot testing activities, and identifies possible directions for future improvements. Annexes provide details on the design and application of the StarPIN methodology: an illustration of how COFOG and CPA classifications can be aligned (Annex 1), a user manual for web-scraping data collection (Annex 2), a set of keywords used to identify the level of innovativeness of services (Annex 3), examples of data sources used to validate data collected through web-scraping (Annex 4), and an illustration of procedures used to upscale website data from the city level to the regional and national levels (Annex 5).

2

The StarPIN approach

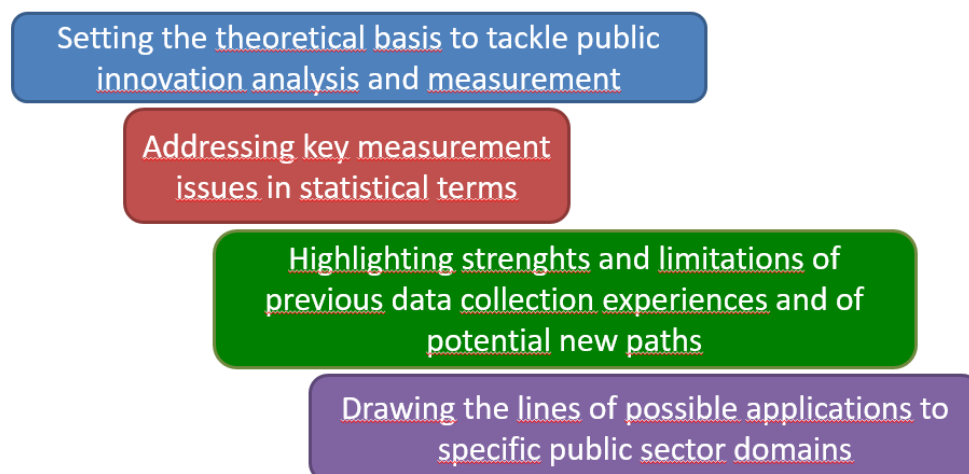
This chapter presents the main pillars of the StarPIN approach to the analysis and measurement of public innovation.

The project is based on the construction of a measurement framework which has taken into account developments in extant literature and data sources, the results of pilot applications of the methodology, as well as the discussions the StarPIN team had the opportunity to have with Eurostat.

The StarPIN project has been characterised by the logical structure illustrated in the diagram below. As shown, based on a critical assessment of the literature on public innovation some key concepts deemed relevant for the measurement exercise have been identified; then, strengths and limitations of extant data sources have been highlighted; thereafter the lines of new and promising paths of data collection (StarPIN methodology) have been sketched. These methodological lines are then applied to explore innovation in specific public sector domains. Based on previous experience in this field of the StarPIN team, the analytical framework has been refined and possible avenues for the improvement of the StarPIN methodology itself have been identified.

Figure 1. Main logical steps of the project

Main logical steps of the project



The conceptual pillars of the StarPIN approach

The StarPIN approach relies on three conceptual pillars: 1. placing public service innovation at centre stage, 2. adopting an object-based approach, and 3. interpreting public innovation as an increase in public value. These conceptual pillars are based on **two largely complementary strands of literature** (discussed in detail in Chapter 3). The first stream of contributions — focusing on **public service dominant logic** — draws the attention of scholars and practitioners to the fact that services are the most important dimension of public sector activities. Hence, it is this area that the analysis of public innovation must focus upon. Emphasising this dimension of public sector activities is consistent with the widespread perception that service provision is a fundamental mandate of government at all levels, even in an era of shrinking public budgets. Moreover, the key role of public services also reflects a general trend that can be observed towards the ‘servitisation’ of the economy as a whole. The second strand of literature — the one on **public value creation and co-creation** — highlights the need for a reflection on the economic and social consequences and outputs of public service development and the identification of aims pursued through public innovation. From this perspective, a focus on public R & D expenditure and technology adoption makes sense only if associated with a consideration of actual usage/effects in terms of new or increased value created for society, either considered as a whole or in its components.

Placing public service innovation at centre stage

The starting point to measure public innovation can be services as key locus of value creation and co-creation. It is suggested that service innovation — as distinguished from mere service delivery (which is per se associated to value creation and co-creation) — needs to be associated with an ‘increase of value’ created and co-created for public purposes. From this perspective, not all value creation can be considered innovation, as long as no change in value creation is determined at some stage of development, provision or usage of a given service. Even the introduction of a novel service cannot be considered public service innovation per se, to the extent that the corresponding change in the value created for society is nil or negative. Hence, the (radical or incremental) novelty of a service is a necessary but not sufficient condition for the identification of public service innovation.

Adopting an object-based approach

Addressing public service innovation implies that a large number of heterogeneous (public and private) players and of interactions among them is to be considered, hence making it particularly hard to identify a single (or a few) ‘representative subjects’ as units of analysis. See Chapter 5 and Annex 1 for a systematic consideration of the variety of such subjects when considered across different government functions and across countries. This extreme heterogeneity induces to explore alternative avenues, other than a pure subject-based approach. A case has to be made here for an ‘object-based approach’ to public service innovation implying a greater emphasis on changes in service characteristics and in the objectives they pursue. This implies reversing the usual approach: from the identification of subjects as the key unit of observation leading to classify innovative activities correspondingly to the identification of innovative objects leading to the identification of subjects involved in the generation and diffusion of such innovative objects. This change in perspective poses important challenges to measurement and statistics. A possible way to tackle such challenges might be that of singling out a manageable variety of public services and of service

characteristics, in a way that is as consistent as possible with current classifications in official statistics.

From this perspective, public service domains wherein innovation takes place can be circumscribed consistent with standard statistical classifications, allowing a link of service activity codes (derived from CPA classification) with COFOG government functions and with NACE codes to identify the set of subjects involved in innovative activities.

The boxes below provide short descriptions of the CPA, COFOG and NACE classifications. See Chapter 6 below for a specification of how this conceptual discussion translates into a data collection methodology.

The statistical classification of products by activity ⁽¹⁾, abbreviated as CPA

The classification of products (goods as well as services) at the level of the European Union (EU).

Product classifications are designed to categorise products that have common characteristics. They provide the basis for collecting and calculating statistics on the production, distributive trade, consumption, international trade and transport of such products.

CPA product categories are related to activities as defined by the statistical classification of economic activities in the European Community (NACE). Each CPA product — whether a transportable or non-transportable good or a service — is assigned to one single NACE activity. This linkage to NACE activities gives the CPA a structure parallel to that of NACE at all levels.

The classification of the functions of government ⁽²⁾, abbreviated as COFOG

The classification was developed in its current version in 1999 by the Organisation for Economic Co-operation and Development and published by the United Nations Statistical Division as a standard classifying the purposes of government activities.

The classification has three levels of detail:

Divisions;

Groups;

Classes.

Divisions describe the broad objectives of government, while groups and classes both define the means by which these broad objectives are achieved.

(1) [https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_products_by_activity_\(CPA\)](https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_products_by_activity_(CPA))

(2) [https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_economic_activities_in_the_European_Community_\(NACE\)](https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_economic_activities_in_the_European_Community_(NACE))

The statistical classification of economic activities in the European Community ⁽³⁾, abbreviated as NACE

NACE is the classification of economic activities in the European Union (EU); the term NACE is derived from the French *Nomenclature statistique des activités économiques dans la Communauté européenne*. Various NACE versions have been developed since 1970.

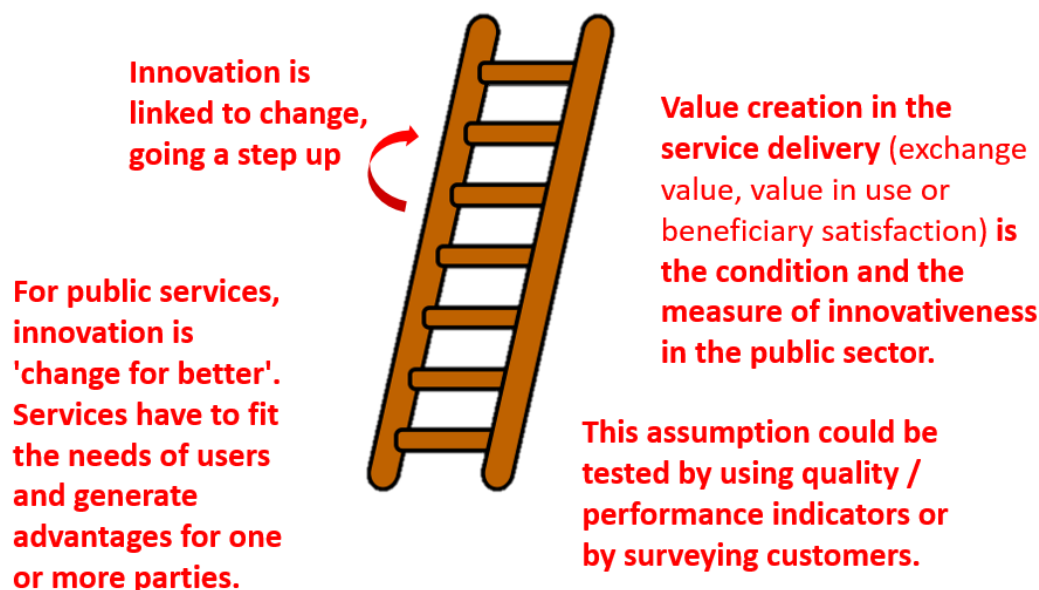
NACE is a four-digit classification providing the framework for collecting and presenting a large range of statistical data according to economic activity in the fields of economic statistics (e.g. production, employment and national accounts) and in other statistical domains developed within the European statistical system (ESS).

NACE Rev. 2, a revised classification, was adopted at the end of 2006 and, in 2007, its implementation began. **The first reference year for NACE Rev. 2 compatible statistics is 2008, after which NACE Rev. 2 will be consistently applied to all relevant statistical domains.**

Interpreting innovation as an increase in value creation

Value creation involves economic aspects (cost reduction, efficiency, a higher quantity and quality of public sector outputs), social aspects (inclusion, quality, and fairness), political aspects (electoral targets) and relational aspects (trust, empathy). One should thus distinguish between innovation patterns according to their contribution to increasing the value creation effects, and capture whether the effect is on the entire communities or on individual members; and whether the effect is direct or indirect. This might be particularly important when conducting surveys that may usefully complement the usage of other data sources and statistics in this field. However, other sources of information can also capture part of the value creation process. Web-scraping procedures can play a fundamental role in detecting the existence and diffusion of service characteristics that can be associated to different levels of innovativeness. Actual and perspective characteristics of services can identify different innovativeness levels symbolised as 'rungs' in the innovativeness 'ladder'. Service providers can be compared with one another in terms of the 'rung' they occupy in a given service ladder at a given moment in time. Changes in innovativeness can be described as movements from one rung to another over time along ladders, with the extension of existing ladders to include non-pre-existing rungs, or even with emergence of non-pre-existing ladders, see the figure on the next page. Such changes can be detected as the appearance of a service characteristic at time t that could not be detected at time $t-1$, or as the emergence of entirely new services.

⁽³⁾ [https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_economic_activities_in_the_European_Community_\(NACE\)](https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_economic_activities_in_the_European_Community_(NACE))

Figure 2: The innovativeness + value ladder

21

As already mentioned, value creation involves economic aspects (cost reduction, efficiency, a higher quantity and quality of public sector outputs), social aspects (inclusion, quality, fairness), political aspects (electoral targets) and relational aspects (trust, empathy). All of these aspects of value creation can be assessed in terms of the component of society being affected; and of the direct or indirect effect that is being produced on such components of society, via the introduction and development of a new or improved service.

The most apparent value creation effect of service innovation is its **direct effect** on the higher quality and/or lower cost of the service being offered. As many public services are not priced, the measurement of such direct effects primarily requires a complex assessment of the quality perceived by users. If a public service can be 'priced', a decrease in prices might reveal an increase in value to the extent that there is no decrease in perceived quality. Service innovation may also increase public value by directly reducing the costs sustained by users in the co-creation of the service itself. For instance the introduction and development of a new electronic certificate may reduce the costs of mobility and of paperwork sustained by the user. Waste collection costs may diminish for firms or citizens as a 'pay-as-you-throw' option is made available. A new or improved service may imply changes in the costs of the service provision, e.g. by reducing or increasing the number of persons or of person/hours involved in its delivery. Strictly speaking, such a cost reduction on the supply-side should not be considered innovative to the extent that it does not also translate into some benefit for society. It remains that a reduction in the cost of providing a given service may in itself be considered a social benefit, if it frees some resources for alternative social uses.

The development of a new or improved service may also have **indirect effects** on value creation, via the reduction of negative externalities or the creation of positive externalities (i.e. not paid for advantage accruing to the society from service provision). This is clearly the case with environmental protection services eventually yielding a reduction in air, land or water pollution; or educational services, increasing the quality of human capital. The development and diffusion of e-government services may reduce the cost for searching for information and monitoring administrative procedures.

Such benefits created via changes in services delivered may well accrue to the **entire community**, in terms of economic outcomes, and/or socially valuable results. To illustrate, new or improved waste management procedures, by increasing the amount of separate collected waste, may improve the quality of environment with a net benefit to the whole society (even in the presence of extra-costs for

polluters). In other circumstances, new or improved services may benefit **some components of society** more than others may. This is particularly the case with regard to services that are more 'vertical' than others (i.e. they apply to specific user categories, rather than being targeted more horizontally to a broader range of applications and uses), as in the case of specific healthcare services for the elderly. It may also occur that a new or improved service negatively affects public value — e.g. when a new mode of service provision is less efficient or effective than a previous one — or generates benefits for one component of society at the expenses of other components. In the latter case, one should evaluate the net social benefit, and it would be correct to consider a service innovative only if it increases the public value created net of social costs sustained by part of the society. This is for instance the case with a new environmental protection service that increases compliance costs for polluters, with net benefits for the rest of society that are greater than the increased compliance costs.

To summarise, one should distinguish between innovation patterns according to their contribution to increasing the value creation effect, and capture whether the effect is on the entire communities or on individual members; and whether the effect is direct or indirect. This might be particularly important when conducting surveys that may usefully complement the usage of other data sources and statistics in this field. However, as will be discussed in section 1.5, other sources of information can also capture part of the value creation process. See Table 1 below for an example of how the effects of public services could be classified.

Table 1: Matrix on public value effects of service innovation

| Effects | Direct effects | Indirect effects |
|-------------------------|--|---|
| Beneficiaries | | |
| Entire community | Lower costs for citizens with the development of pay-as-you-throw schemes Greater transparency of administrative procedures and lower costs of mobility for users of new e-certificates | Water quality improvement as a result of better waste management More skilled workers as a result of better educational services |
| Specific groups | Increase in the number of patients assisted via homecare health services | Reduction of infection rate diffusion through new vaccine campaign |

A new hierarchy of data sources

The StarPIN approach thus implies that public services (objects) are the primary focus for the analysis of public innovation, and that relevant innovators (subjects) are identified in accordance with existing statistical classifications. Data need to be collected from different sources to capture the nature and intensity of innovation.

New data sourcing opportunities (from social networks to sensors, from the web to open data, from a direct participation of individuals to generate new data to the 'aggressive' ability of private large companies to appropriate enormous quantities of data and extract big value from them) have key implications for data collection and management. This 'explosion' of opportunities generates the need for greater abilities to combine a variety of data sources and to select among huge quantities of data.

A new hierarchy of data sources is suggested, with web-scraping providing a promising starting point for data collection in this field. Other sources of relevant information on the characteristics and performance of services under observation, including administrative data, can be used to validate the results of data obtained from websites, enrich our understanding of innovation in any given public sector domain and geographic area.

Chapters 6 and 7 in this working paper illustrate the main data sources that can be utilised in the case of two specific public service domains in five EU countries, with a particular emphasis on web-scraping procedures and administrative data.

Eurostat has conducted in recent years a project linked to increased use of administrative data. Eurostat's VIP ADMIN Project ⁽⁴⁾ and its outcomes document improved collaboration between public entities to make better use of available administrative data for official statistics. The wide range of administrative registers available in the Member States have evolved into coherent and updated tools due to newer informatics applications and ICT skills of users. The following reports of Eurostat's VIP ADMIN Project ⁽⁵⁾ are now final and publicly available: *Good practices in accessing, using and contributing to the management of administrative data*, and *Analysis of the legal and institutional environment in the EU Member States and EFTA Countries*.

It is argued that these data sources help fill the gap in existing official statistics and provide highly complementary pieces of information, shedding light on innovative phenomena in the examined public service domains. Also ad hoc surveys could be consistent with the StarPIN approach, to the extent that the population of innovators (subjects) to be monitored is designed in a way that is functional to exploring specific public service domains (objects), and not vice-versa. Moreover, in Chapter 6 below it is suggested that some important complementarities exist also between web-scraping data-mining and the design of ad hoc surveys.

The choice of the object approach also has consequences for the use of particular methods of site-centric measurements, implying a shift from syntax-based methods towards semantic-based methods. The former takes the readable text of an entire website (of particular technical components or modules) as the unit of observation. The unit of analysis is then the website as a whole, which refers to the website owner or the organisation. Since one public organisation often provides several public services on one website, and these services share a common technical platform (i.e. certain generic components such as authentication and payment modules) it is difficult to develop a generic model to map particular pieces of the source code to particular public services. **The only way to do so, and hence adopt individual services as the main focus of analysis, is via semantic-based methods.** These methods take the actual content of the webpages (as shown to the visitor of the website, thus after rendering by the web browser) as units of observation (Bianchi et al., 2017). In the specific case of StarPIN, these are the descriptions of public services at specific web pages of the website of a public organisation. It is important to point out that it is not measured the public service itself but rather the description of the service. From this description the relative level of innovation achieved by the institution introducing the service concerned can be inferred by adopting a comparative method. As a consequence, both a predefined list of innovation levels and a specific set of keywords for each level of innovation of a single public service have to be made available to analysts. This is the semantic (content) based, ex ante (top down) measurement of topics (Schut and te Velde, 2017). See Chapter 6 below for a discussion of how keywords are identified and used to describe service characteristics, Chapter 7 for an application of the keyword-based methodology in specific public sector domains, and Chapter 8 for lessons learned from practical use of this methodology.

⁽⁴⁾ https://ec.europa.eu/eurostat/cros/content/2016-workshop-access-administrative-sources_en more information at <https://ec.europa.eu/eurostat/web/ess/admin/newsletters>

⁽⁵⁾ The reports are available on CROS: https://ec.europa.eu/eurostat/cros/content/wp1-access-administrative-data_en

Using (semantic-based) web scraping to collect data on public sector innovation implies the definition of criteria to classify service characteristics as 'potentially innovative'. Experts in the field can be used to provide guidance in the description of services and in singling out relevant data sources. Once designed with the support of experts, further refinements and replication over time of data collection should be possible within the general guidelines set at the beginning of the process, with no further expert control. The method proposed by StarPIN thus introduces the need to devote more substantial resources to the preparation phase rather than to the mere data collection phase. This brings as a consequence a dramatic drop of costs in replicating the measurement year by year or even changing the measurement frequency.

3

The foundations of the StarPIN approach

This chapter discusses the theoretical foundations of the StarPIN approach by illustrating and combining two main streams of literature: contributions on the public service dominant logic and studies on public value creation and co-creation. The former stream of contributions — focusing on **public service dominant logic** — draws the attention of scholars and practitioners to the fact that services are the most important dimension of public sector activities; hence, it is this area that the analysis of public innovation must focus upon. Emphasising this dimension of public sector activities is consistent with the widespread perception that service provision is a fundamental mandate of government at all levels, even in an era of shrinking public budgets. Moreover, the key role of public services also reflects a general trend that can be observed towards the ‘servitisation’⁽⁶⁾ of the economy as a whole.

The latter strand of literature — the one on **public value creation and co-creation** — highlights the need for a reflection on the economic and social consequences and outputs⁽⁷⁾, of public service development and the identification of aims pursued through public innovation. From this perspective, a focus on public R & D expenditure and technology adoption makes sense only if associated with a consideration of actual usage/effects in terms of new or increased value created for the society, either considered as a whole or in its components.

The chapter will first place scholarly research on public innovation in a more general context, and then discuss in greater detail the two main streams of literature we have just mentioned.

Combining insights from complementary strands of literature

Public innovation has been addressed by a rather extensive number of very heterogeneous contributions encompassing different disciplines, objects and approaches. This literature ranges from analyses of regulatory changes, to studies on the evolution of governance modes, on the diffusion of new technology within public administration, and on new service adoption and development. Nevertheless, concentrating our attention on public service innovations, which are at the core of our analysis, makes our effort to review the literature more focused and manageable. When dealing with public service innovation, the starting point is the acknowledgement in extant literature that most

⁽⁶⁾ The increased importance and pervasiveness of the service sector in industrialised economies is referred to as the servitisation of products or a product-service system. Virtually every product today has a service component to it.

⁽⁷⁾ The distinction between innovation output and innovation outcome has been discussed in the literature in a number of occasions. E.g. for the private sector, ‘Innovation outcomes are the consequences of the introduction of innovations, among them the economic effects of innovation outputs’; or ‘... innovation output may be indicated by statistics on patents, papers, prizes, invention disclosures, and degrees awarded, while outcome proxies include patent and paper citations, expert evaluations, innovation counts, new product sales, measured productivity growth, and benefit/cost or rate-of-return estimates.’ At this stage of the work, a choice was made of not distinguishing between the two terms/concepts.

innovation literature has developed with reference to technical change in manufacturing processes and has long disregarded innovation in services. Even more importantly, if innovation in services is the Cinderella of innovation literature, innovation in public services is the Cinderella of the literature on service innovation (Miles, 1998).

Various differences and peculiarities were highlighted in the economic literature when moving the focus of analysis of innovation processes and results from production and industrial products in manufacturing sectors to the service sectors ⁽⁸⁾. An even more relevant set of observations relates to the distinction between the dynamics concerning the private service sectors and the public service composite and diversified world. A synthetic picture of similarities and divergences is provided in the following box:

Similarities

- Various phases and components of the innovation process which are both distinct and strongly interconnected;
- Non linearity of innovation process;
- Both promoter and user are involved and play key roles;
- Contextual factors condition innovation all along its development.

Divergences

- No market;
- Price factor is not a quality indicator;
- Objectives and strategies significantly differ from the ones in business sectors.

It is difficult to provide for a systematic and comprehensive analysis of the multifaceted issue of service innovation but some specific issues are particularly relevant for extending the analysis to the public sector and will be discussed in the next paragraphs.

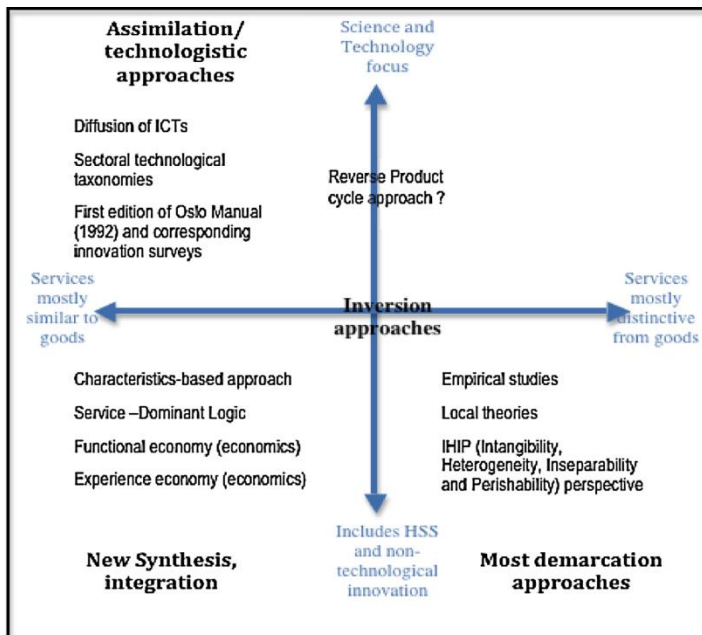
Approaches to public service innovation

The starting point for our review of public innovation is the taxonomy of approaches proposed in Djellal et al. (2013). This contribution focuses on public service innovation and draws attention to four different theoretical perspectives that have largely emerged in the literature on business services, but with applications and extensions that may be useful for the study of public services innovation ⁽⁹⁾.

1. **Assimilation** approaches, focusing on the adoption of new technologies in shaping innovation in services;
2. **Demarcation or differentiation** perspectives, which emphasise that service sector evolution may follow heterogeneous patterns largely independent from technical change occurring in manufacturing industries;
3. **Inversion** approaches, which see (some) service industries as being sources of innovation across the whole economy, mobilising their suppliers, or providing inputs to innovation among their clients;
4. **Integrative or synthesis** perspective, which 'seeks to provide the same analytical frameworks for both goods and services products, for manufacturing and service industries, and for both technological and non-technological forms of innovation' (Djellal et al., 2013).

⁽⁸⁾ See the differentiation approach recalled in the following paragraph.

⁽⁹⁾ Reference studies cited in Djellal et al. (2013) for these four perspectives are Gallouj, 1998; Gallouj and Weinstein, 1997; Coombs and Miles, 2000; Droege et al., 2009, and other authors cited in the latter study.

Figure 3: Taxonomy of approaches for evaluating public service innovation

source: Djellal et al., 2013

Going into a more detailed review of the four perspectives, it is worth noting that the ‘**assimilation approach**’ has long prevailed in the literature. Some studies which can be considered as part of this perspective have contributed to monitoring service evolution and service innovation, identifying concepts and analytical tools that are largely applicable to public services. This is the case with the wide set of studies on the contribution of information technologies (IT) and information and communications technologies (ICT) and, even more, the development of networking and mobile and ubiquitous IT and social networks, on public sector organisation, public service content and potential, interaction and participation of beneficiaries. As outlined by Djellal et al. (2013), this group of approaches includes: empirical work devoted to analysing the diffusion of exogenously developed technology into services; the reverse cycle model of Barras (1986, 1990) and its applications to public services especially at local level, the learning process both in service process and in product along the technology waves, up to the online delivery of public services considered as a product innovation, e-government, dematerialisation, e-democracy, online reimbursement, and interaction between administrations and users. Still this approach emphasises technological innovation in particular⁽¹⁰⁾ together with sectoral taxonomies that trace distinctive modes of innovation in various service industries, i.e. service innovation patterns. Several public service entities can be put in the supplier-dominated category, largely adopting IT technologies taken from sources that are external to the public sector. But other ones are similar to large private network service organisations and because of their size and relevance are objects of interest for dedicated technology applications. Different public services rely on different knowledge bases, and relate to citizens in quite different ways. R & D departments and programmes, laboratories and service design efforts can be observed in many public service sectors. Various technological and non-technological trajectories can coexist separately or in hybrid fashions, in private or public service organisations (Djellal and Gallouj, 2005, 2008). ‘Different innovation patterns and trajectories can coexist and interact within the same organisations.’ (Djellal et al. 2013, p. 105)

Another topic discussed in the literature is the standardisation of services in line with the product standardisation. The industrialisation model for public services is aligned with the ‘new public management’ approach, which looks for efficiency by adopting more market-like elements within

⁽¹⁰⁾ here is little attention to new non-technological services that may be created, such as new social policies, like pension entitlements, citizen rights, ombudsman procedures, (...) new services in support of active independent living that employ new systems for monitoring the health of vulnerable people.’ (Djellal et al., 2013, p. 103)

public administrations, more outsourcing of services, more competitive pressures, targets and performance measurement indices.

Two 'loosenings' of the assimilation approach have been identified in previous reviews of the literature: the autonomisation, which interestingly brings to a complete reversal of the power relationship, generating innovation trajectories that are no longer 'supplier-dominated' but 'consumer or user-dominated' ⁽¹¹⁾, and the endogenisation of IT innovation, which combines tangible and intangible IT and organisational engineering.

Having appreciated the merits of the assimilation perspective and of its extensions, one could argue that this approach also has clear limitations when dealing with many aspects of public service innovation. There are several reasons to criticise this approach, which has long dominated research in this field:

- i. An important part of public service innovation does not rely on the adoption of new technology per se, but often has to do with the creative implementation of existing technology and with its usage to improve existing services and to design and develop new ones;
- ii. By emphasising the innovative function of technology suppliers that are external to the service sector, and of service providers as adopters, the assimilation approach largely disregards the role of users in service innovation;
- iii. Being designed to interpret service innovation in the private sector, the approach has a hard time at addressing the innovative behaviour of not-for-profit organisations and of public sector entities that are characterised by a multiplicity of aims;
- iv. If applied to the public sector, the assimilation approach underplays the role of the public sector itself in intermediating between final users and suppliers of technology, and eventually acting as a catalyser of innovation that occurs elsewhere from the public sector (Edquist and Hommen, 2000);
- v. Altogether, the approach leads to underestimating the heterogeneity of actors involved in service innovation and of public service innovation in particular.

Being aware of these limitations of the assimilation approach, special attention has been devoted to other approaches (including contributions that are at the crossroads of different perspectives as in the case of reverse product cycle literature):

Let us consider the so-called **demarcation or differentiation** approaches, which assume that service innovation will follow specific patterns that may be largely independent of the evolution of manufacturing. This emphasis on differentiation factors marks a departure from the central idea in the assimilation approach that sources of heterogeneity will eventually come from outside the public sector.

Criteria and characteristics highlighted by the literature grouped under this approach by Djellal et al. (2013) are: distinction between back office and front office; intangibility and interactivity where the role of the user as contributor in producing the service emerges, co-production; non-market stimuli/motivations to innovation; the influence of the political system; the principles of public service i.e. equality, fairness, continuity. Along these issues, service activities are examined in knowledge processing and IT applications, and, with a demarcation emphasis, specific forms of innovations emerged: administrative innovation, systemic innovation, conceptual innovation, radical change of rationality, positioning, strategy, governance and rhetoric, institutional innovation.

The demarcation approach mainly generates qualitative empirical research. The Oslo Manual, that included marketing and organisation innovations and the Community Innovation Surveys (CIS), has been extended to the service industry and to some non-technological forms of innovation, but still not fully covered are: 'non-technological product innovations, non-technological process innovations, ad hoc innovations and tailor-made innovations, innovations in complex "packages" of goods and services, also known as new concepts or new formulas (for example, in retail, hospitality, etc.), social innovation, user innovation by consumers, i.e. innovations introduced by consumers who develop or modify goods and services for their own use' (Gault, 2012, p. 109). Indicators developed for subsectors of public services, e-health and local government, and broader indicators developed at the country level, with Mepin at the EU level, the Innobarometer 2010 attempt was dedicated to

⁽¹¹⁾ Which opens to the analysis of public procurement role in guiding and stimulating private sector innovation.

innovation in the public sector. This approach is associated with professional or cognitive rationalisation, as opposed to industrial rationalisation, 'divided into three strategies: typification of cases, formalisation of problem-solving procedures (methods), and the use of individual or organisational routines' (p. 9), from mass (long dominant) to personalised services.

As for '**inversion** approaches', reviewed literature emphasises the role of (some) service industries as being sources of innovation across the whole economy, mobilising their suppliers, or providing inputs to innovation among their clients.

This approach provides the framework for the analysis of potential impact of service innovation (one of the critical ones being knowledge-intensive business services — KIBS) on specific sector activities at both the microeconomic and macroeconomic level. Studies included under this heading identify service innovation as corresponding to a sort of 'Schumpeter Mark III model' wherein the key engine of innovation is the interaction between different sectors of the economy (Gallouj, 2002). There are four possible levels of impact of public service innovation on other sectors: (a) public services (e.g. public R & D, and technology transfer services) can support innovation in other economic activities; (b) public service providers can be lead users of technology, thus stimulating innovation of suppliers (via public procurement); (c) public services may be responsible for providing valuable information to private sector; (d) public service providers can be co-producers of innovation, or involved in Public–Private Innovation partnerships. There are connections here with the integrative or synthesis perspective.

As far as the '**integrative or synthesis** perspective' is concerned, the main effort is to 'develop unified analytical frameworks which can accommodate relevant aspects of innovation for both goods and services, for manufacturing and service industries, and for technological and non-technological forms of innovation' (Djellal et al., 2013, p. 99). Emphasis is given to the shift from a linear model to an open model of innovation that appears to characterise both services and manufacturing activities, blurring the boundaries between sectors and leading to the 'servitisation' of the whole economy (see Vandermerwe and Rada, 1988; Howells, 2001; Neely, 2008 cited in Djellal et al., 2013, p. 111).

Contributions along this direction of research emphasise that the actual value of most goods is by and large determined by their information and service content. The emphasis is on (private and public) providers of solutions, functions or experiences that can add value to products. New theoretical research prospects have emerged, including functional economics (Stahel, 1997), experience economics (Pine and Gilmore, 1999), the service-dominant logic (Vargo and Lusch, 2006), and characteristic-based approaches (Gallouj and Weinstein, 1997; Windrum and Garcia-Goni, 2008; De Vries, 2006; Gallouj and Toivonen, 2011, Savona and Steinmueller, 2013). In Djellal et al.'s (2013) view, all of these streams of literature can contribute to address innovation issues (see the building blocks of the public service innovations proposed in Ch. 2.3).

As for the characteristic-based approaches, 'a service can be defined as the mobilisation of internal or external competences and internal or external techniques (tangible or intangible) to produce final or service characteristics (that is to say use values)' (Djellal et al., p. 111). Changes in service characteristics or in their combinations result in different models of innovation. Some of these innovation processes can also take the form of invisible or hidden innovation (NESTA 2011), often defined as bottom-up modes (Fuglsang, 2008, 2010). These include unplanned and informal innovation patterns, such as ad hoc model, practice-based model, bricolage model, and rapid application model. It is likely that much innovation of this type occurs at local levels in public service organisations, and one problem may well be that much of this is not widely known and diffused in the organisations (Djellal et al., 2013, p. 113). Other forms of public service innovations are labelled as 'collaborative' (Sørensen and Torfing, 2012), wherein collaboration is envisaged as taking place inside public administrations, between members of different public bodies, but also with private partners. New innovation networks, labelled public–private innovation networks in services (PPINS) have been studied under the European project ServPPIN (The Contribution of Public and Private Services to European Growth and Welfare, and the Role of Public-Private Innovation Networks, FP7).

Key streams of literature on public service innovation

Based on this overview of economic perspectives on public service innovation, the study will focus on the two main streams of literature identified earlier:

The literature on public service dominant logic;
The literature on value creation and co-creation.

These strands of contributions are of particular interest for the purposes of this study as they help to tackle two key issues in public innovation. The public service dominant literature provides some useful insights on the main **object** of analysis that needs to be focused upon when dealing with public innovation that is public services. The value creation and co-creation perspective offers a set of concepts and analytical tools for the analysis of **process** through which public service innovation is generated and on the actual **output** of such generation process.

The literature on public service dominant logic

The idea of **public service dominant logic** put forth by Bason (2012) and Osborne (2012) is that services play the most central role in the public sector, hence leading to abandon a product-dominant approach that has prevailed in most economics of innovation literature. The basic unit of analysis when dealing with public innovation should thus be public services.

The service dominant logic elaborated by Vargo and Lusch (2006), is based on the proposition that organisations, markets, and society are fundamentally concerned with exchange of services — applications of competences (knowledge and skills) for the benefit of a party. From this perspective, services are basically exchanged against other services; all firms end up being service firms; all markets are centred on the exchange of services, and all economies and societies are service based. Service dominant logic embraces concepts of the value-in-use and co-creation of value rather than the value-in-exchange and embedded-value concepts that are central in goods dominant logic.

Some of the fundamental premises were enriched or re-formulated in Vargo and Lusch (2008): 'Service is the fundamental basis of exchange, (...) The customer is always a co-creator of value. The enterprise cannot deliver value, but only offer value propositions. A service-centered view is inherently customer oriented and relational, all social and economic actors are resource integrators and value is always uniquely and phenomenologically determined by the beneficiary' (p. 7) Vargo and Lusch (2008).

The service dominant logic extended to the public sector is analysed by Osborne, Radnor and Nasi (2013) looking at the intersection of value creation and value co-creation.

Osborne, Radnor and Nasi (2013) argue that 'public management theory is not fit for the purpose, because it focuses on intraorganisational processes rather than on the interorganisational ones which are the reality of public service delivery, and is too much based on the manufacturing sector managerial theory. This is why the work argues for a 'public service dominant' approach, drawing upon service-dominant theory. This in the era of the New Public Governance'.

'The New Public Governance', as the new paradigm of public administration science, emphasised pluralism, attached great importance to the links between internal and external organisations, and paid attention to organisational governance. 'The New Public Governance' paradigm had different names, such as 'the New Public Governance' (Osborne, 2006), 'New Governance' (Rhodes, 1996), and 'Public Governance' (Skelcher, 2005), etc., cited in XU, SUN and SI (2015).

Bason (2010) emphasised 'the four Cs' in the public sector innovation and the ability to create ideas, implement them and create value for citizens and society: Consciousness, Capacity, Co-creation and Courage. These dimensions are strongly rooted in the subject/process side of public service innovation, deeply based on a concrete managerial experience in the public sector enhanced with a strong academic refinement of the practical work. The combination and reciprocal support of these

two components of Bason's experience brought him to elaborate a human-centred governance and to attribute/recognise a critical role to design in public service innovation. (Bason, 2017)

A possible extension could be to focus on the so-called '**General-interest services**', i.e. services considered to be in the general interest by the public authorities and accordingly subjected to specific public service obligations. They include non-market services (e.g. compulsory education, social protection), obligations of the state (e.g. security and justice) and services of general economic interest (e.g. transport, energy and communications) (EU-Glossary).

It would be misleading to consider such activities as 'products' or even 'public goods'. In fact public services are largely (albeit not exclusively) intangible in nature, are characterised by important complexities of innovation processes, and by a wide variety of the organisational resources and skills involved. As suggested by Osborne, Radnor and Nasi (2013, p. 149), adopting a public service-dominant approach has four important implications: 1. The acknowledgement that both the citizen and user are essential to public service delivery processes and their engagement in these processes adds value to both; 2. Public sector organisations are increasingly involved in shaping the expectations by their users hence affecting the role of the latter in innovation activities; 3. Coproduction becomes an inalienable component of public services delivery; 4. New managerial skills and organisational capabilities are needed to make service provision effective.

Design emerges as one of the approaches which 'can contribute in powerful ways to innovation in the public sector and, potentially, transform how public organisations are governed' (Bason, 2017). Public service design can be one of the areas of activity and of cooperation between public service providers and beneficiaries/users bridging the two streams of literature, as it stems from the public service specific domain but it can be also analysed for its co-creation content.

The literature on 'value creation and co-creation'

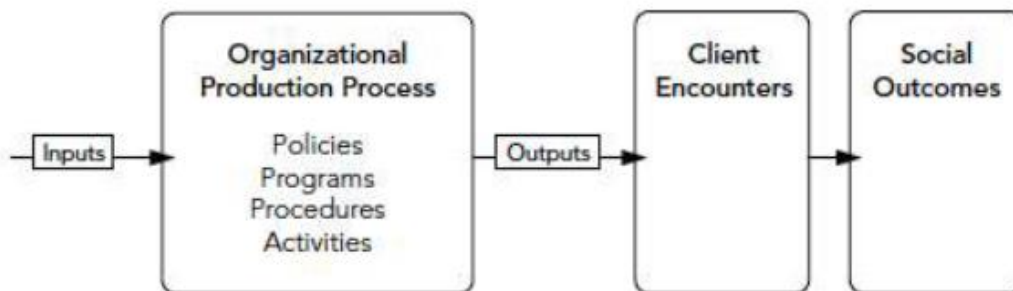
This literature tackles the key issue of innovation output, and of the process of its generation. An emphasis on value creation and co-creation is present in both business service innovation and public service innovation contributions. However, value creation, considered in a wide sense, poses relatively less conceptual and measurement problems in the case of the private sector than in the case of public sector. In fact, in the private sector product and process innovation may be thought of as leading to greater value in terms of additional turnover, revenues, and profits; while changes in organisational practices can reasonably lead to increases in productivity. Shifting to the analysis of value creation in the public sector poses serious challenges in this respect. Reference will be made to attempts to measure value creation via performance indicators in the public sectors (see inter alia the review on ways to measure performance in the Public Administration conducted in Seri and Zanfei, 2013).

A wider concept of public value is not strictly related to the public sector performance but rather refers to values benefitting the community or values generated by the reduction of negative externalities.

Moore (1995) contends that in the public sector the arbiter of value is not the individual, but the community — that is, society in general, acting he says 'through the instrumentality of representative government' — and likely to be made up of service users, tax payers, service providers, elected officials, treasury and media. The issue of who is entitled to evaluate public value creation is left unresolved.

Moore (2012) — building on his book 'Creating Public Value' — proposes a value chain for public administration, and indicates possible points of measurement in the following scheme) (see Figure 4):

Figure 4: Value chain of the public administration (source: Moore, 2012)



The development and use of some kind of accounting scheme is needed, in order to capture — or, in the accounting sense, recognise — when significant costs to the society were being incurred, and substantial benefits in the society being produced. (Moore, 2012, p. 2)

There have been several attempts to develop such a scheme based on a variety of philosophical approaches ⁽¹²⁾.

Moore (2013), proposed a different path for future development of measures with the aims of: '1. reliably capturing the public value being produced by a government enterprise; 2. meeting public demands for accountability in government; and 3. helping managers to use performance measurement systems to guide their organisations towards improved performance' (p. 11)

This path is based on various elements:

Taking a Developmental, Strategic Approach, which in turn will require: (a) acceptance of the fact that the performance measurement and the value created measurement will change over time, (b) the need for government managers 'to have some way of monitoring their current position and capacities, and to envision and execute strategies that would sustain or improve their position and performance in the future.'

Organising Four Different Kinds of Managerial Work:

- i. Technical, 'the challenge of developing specific measurement instruments that can accurately capture the degree to which valued (and disvalued) effects are occurring in the world while the enterprise is operating' (p. 13)
- ii. Managerial, 'the actions that managers have to take to build the performance measures, and to use them in their organisations to drive performance.' In other words: 'All one has to do is to attach specific performance measures to particular managers or workers, and use them to reward good performance and punish bad, and the performance of the organisation will improve — at least in terms of the measured dimensions of performance.' 'Things get even more complicated if (...) the managers decide that they want to help the organisation learn about what works rather than simply work harder.' (p. 14)
- iii. Philosophical. This type of work is thought of by Moore (2013) as even more relevant, but definitely less studied up to recent time: 'The idea of social or public value is a normative, philosophical concept. (...) it is important to recognise that all performance measurement and management systems have at their core a normative theory of what would be valuable and what would be costly if an organisation produced those valued or disvalued effects.' (p. 14) Moore distinguishes between two philosophical traditions, both to be taken into

⁽¹²⁾ Hatry, H., Wholey, J. and Newcomer, K., eds.(1994), *Handbook of Practical Program Evaluation*, Jossey Bass, San Francisco; Weiss, C. (1973), 'Where Politics and Evaluation Meet,' *Evaluation* 1 (1973): 37-46; Hatry, H., Winnie, R. E. and Fisk, D. M. (1981), *Practical Program Evaluation for State and Local Governments*, Urban Institute Press, Washington D.C.

account while discussing about public value: on one hand, individual utilitarianism and social or collective one and the correspondent material conditions, on the other hand, he refers to the deontological traditions, which use the concepts of duties and responsibilities, along legal or juridical ideas, within a normative language ('fairness, justice, right relationships, duties and responsibilities, rights and privileges') (p. 14)

- iv. Political, '(...) the political work (...) is crucial, of course, because in a democratic society and a democratic government it is not enough to have a system that is philosophically and technically sound, and managerially useful; it also has to be endorsed and embraced by those in positions to authorise, legitimate, and pay for the enterprise. This is **morally** important because the only appropriate arbiters of public value in democratic political systems are citizens, taxpayers and their elected representatives. It is **practically** important since it is the political demands for accountability that create the drive to create and energy to develop and use performance measurement system.' (p. 15)

These four categories of work activities should be kept in mind when **analysing and ranking public services and the aims they answer to**. They can maybe be used as criteria for assessing the emergence of new public services to satisfy new facets of public values.

In this framework the 'accountability agents' play a key role. These are defined as 'social actors who feel that they have at least the right and sometimes the responsibility to call public organisations to account for some aspect of their performance'.

Developing the Public Value Account

Such an account has to cope with specific public sector features:

- i. 'the assets that government uses to produce socially valuable results are not restricted to money. The government can alter material and social conditions in the world by using its **authority** as well as its money.' (Moore, 2013, p. 21)
- ii. '(...) individuals and their satisfaction are not necessarily the appropriate arbiter of public value.'
- iii. '(...) when the government acts, it will be judged by deontological standards of fairness and justice as well as utilitarian standards of satisfying individual clients or achieving desired social outcomes.' (p. 22) 'there are many clients of government agencies.' and additionally some desired aggregate social welfare' which might not correspond to any individual satisfaction. 'legislating or public policy-making' (p. 23)

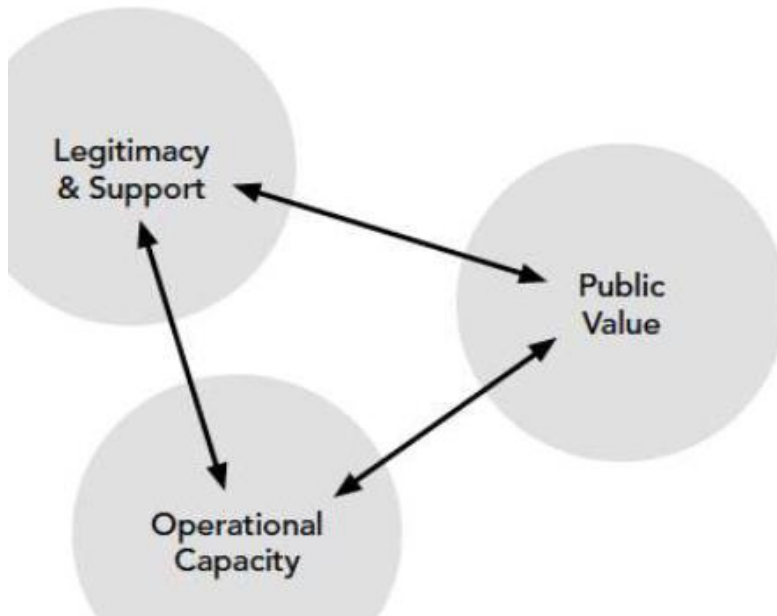
On the outcomes side: '(...) the satisfaction of clients of government agencies as an important value created by the government, (AND) (...) the specific values described as the mission or the ultimate goals of the government enterprises.' (p. 23)

Moore (2013) 'added the paradoxical idea that a public organisation might want to include room in the accounting scheme for recognising unanticipated good and bad effects of its activities that were not part of its original mission.' (p. 24)

Developing a Public Value scorecard, adding 'families of measures that will focus managerial attention on the current position of the organisation in its environment, whether and how operations can be sustained, and how both current operations and future positioning might be improved through particular investments that embodied the continuous learning of the organisation.' (p. 25)

The underlying aim is to 'capture a current understanding of what public value we seek to produce'

Figure 5: The Strategic Triangle for Imagining and Testing Public Value Propositions (source: Moore, 2013, p. 30)



Examples of outputs/outcomes to be measured for NY Policy Dept (Moore, pp. 32-3)

In legitimacy and support perspective:

Mission alignment with values articulated by citizens

Inclusion of neglected values with latent constituencies

Status of key legislative and public policy proposals to support enterprise: Authorisations, Appropriations

In Operational Capacity:

Organisational outputs: Quantity of Outputs, Quality of Outputs (attributes that produce desired results, that increase client satisfaction, that reflect current values in operations)

Based on this general framework, Fuglsang and Rønning (2015) analyse the relation between public service innovation and public value. Values can be defined as measures for beneficial behaviour that guide innovation. Value-tensions in public services include tensions between the political, economic, communal, aesthetic and intellectual values. They point to the concept of intertwined innovation patterns, such as the intertwining of **science-driven and task-driven** innovation, and to how **varied values** guide innovation in public services.

Elias (2016) highlights the contribution in the public service value debate of both Benington and Seddon who 'focus on the need to identify purpose as a prime initial task in service improvement, rather than go down the customer/value route. Identifying purpose appears to have strong resonance with public value.' (p. 9). Elias claims that 'Indeed, public value is often at odds with private value.' (p. 14)

4

Towards an analytical framework: topics, issues and variables

Based on the discussion of key concepts and analytical challenges carried out in Chapter 3, in this chapter **the topics, issues and variables** that are relevant for the measurement of public innovation will be singled out. **This will be an essential starting point for the development of a methodology that will be aimed at integrating:**

- i. current (official) statistics ⁽¹³⁾,
- ii. administrative data, and
- iii. data extracted from the web.

In other words, the use of any of these data sources has to be seen as essential to meet the project's requirements. Of course, data sources will have to be organised according to well defined analytical contexts — labelled here as **topics** — which can be described as **areas of public sector activity wherein innovation can take place, and can hence be observed**. Within such topics/areas of public activity, a set of specific observable characteristics of activities being carried out in such domains could be identified — those can be defined **issues**. When analysing public sector innovation, such issues can be considered as **characteristics of innovative activities**. In our terminology, these issues/characteristics of innovative activities can be analysed by means of a set of **variables or indicators/proxies that make such issues measurable**.

This chapter will define and identify in statistical terms the areas of public sector activity (topics) wherein innovative activities can be observed. Chapter 6 will illustrate alternative data sources and data-collection methods to measure some of the key innovation aspects/issues, by means of some of the variables and proxies that can be constructed. As some specific aspects of public innovation activities might not be captured by means of the data sources mentioned above (current statistics, administrative data and data derived from web-scraping), there might also be a need for ad hoc surveys. These will have to be based on sets of questions administered to more or less extensive samples of actors involved in public sector innovation. See Chapter 7 for a brief discussion of some of the topics/areas of public sector activities and/or public innovation issues that might need to be explored by means of such surveys, **and for a concise description of the main lines along which they could be designed**.

The concepts and analytical tools described as topics, issues and variables will be discussed more in detail below describing how these can be operationalised from a statistical point of view.

Topics (areas of public sector activity wherein innovation can take place). Given the relevance of public services as the key output of public sector (irrespectively of the level of administration taken into consideration) a set of public service domains will have to be identified as a sample to be analysed for the purposes of this statistical reporting. See Chapter 3 for a discussion of criteria for defining relevant subsectors and services developed within such subsectors. An example of service

⁽¹³⁾ More on this in the Annexes, e.g. list of the Eurobase as primal source and link to respective statistical domains.

area that can be identified with reference to current statistical classifications is waste collection services (which in turn are a component of the wider service area of waste management).

As described in detail in Chapter 5, each activity area will be related to the public sector functions and service classifications. To illustrate:

- i. government function: environment protection;
- ii. sub-function: waste management;
- iii. focus on the service or one of its components — e.g. distinguishing between hazardous; and non-hazardous waste — and on its assignment to a government level.

For each topic, specific issues have to be identified.

Issues (ways to characterise innovation activities taking place within each of the examined areas of public sector activity). The key issue to be addressed in our view is the degree of innovativeness characterising the observed area of public sector activity (topic) under consideration. It has to be stressed here the distinction between innovative activities — leading to innovative outcomes — and innovation itself, which is the main object of this statistical reporting. Consistent with our line of argument above, the degree of innovativeness should be captured in terms of the (additional) value that is being created. Two dimensions of value creation have been taken into consideration.

- i. Dimension a. — in which way value is created?
- ii. Dimension b. — has the value created increased?

Within each area of public sector activity, the degree of innovativeness needs to be captured by means of:

Variables — (indicators of the presence of a certain degree of innovativeness of the public sector activity under consideration). Our methodology relies on both direct measures of innovation — e.g. the introduction of a service that is new, not supplied before, and can be associated with evidence of the creation of some greater public value — and more indirect, albeit often easier to detect, measures of innovation. The latter include characteristic of an existing service that can be considered as more innovative than another service. To illustrate, the presence of a given service characteristic, such as 'pay-as-you-throw' scheme in municipal solid waste collection could be considered as more innovative than standard waste collection methods. The various forms of providing and delivering services can be sorted out in terms of the potential contribution to increase the value created. The 'change' in value creation can be identified as a shift from one rung of a 'ladder of innovativeness' to another one, or from the combination of more than one rung. See Chapter 2, Figure 2 for an illustration of this logical procedure. The counting of the rungs generates the statistical indicator.

Of course, a variable that captures public service innovation can be either a dichotomous variable, or an ordinal categorical variable that takes as many values as the rungs of the ladder (and/or its combinations).

The ladder illustrates the degree of observed innovativeness, with the criterion that the higher rung corresponds to the most innovative level of innovativeness of that service. The absolute and relative frequency of a given rung can be used as a measure of innovativeness.

Exceptions will include having a categorical variable without pre-ordered value or not having an a priori classification at all, i.e. not being able to identify the direction of the scale.

A variable which captures the dimension b, is: how differentiated are the frequencies collected for the rungs and their combinations. Dimension b can also be bi-variate (innovative/not innovative).

5

A classification grid

This chapter operationalises the conceptual elements sketched in Chapter 2, by tackling the methodological and statistical challenges arising when an object-based approach was adopted.

Having identified public services (objects) and having derived relevant innovators (subjects) consistently with existing statistical classifications, data need to be collected from different sources to capture the nature and intensity of innovation within the public service domains under observation.

In a context where a lot of qualitative evidence on plans and projects for introducing innovations in public institutions is available, an issue should be raised about the key requirements to transform information into statistics which have the level of quality requested by advanced statistical systems like the European Statistical System (ESS).

This is very much an issue of 'relevance', an essential component of the quality check for official statistics in the EU. In this respect, relevance can be interpreted in a twofold sense. On the one hand, statistics on innovation in the public sector — or better, in public services — will have to be relevant for the EU policy. This means using the concepts and definitions developed by the EU institutions and answering to a set of questions still open in the European political arena. On the other hand, a statistical relevance will have to be assured by allowing for the adoption of the standard classifications used in official statistics, mostly of national accounts concepts and definitions (according to the ESA version of them).

The identification of the target statistical unit: public institutions

Several experiments have been designed in order to replicate in the public sector the successful experience of undertaking a direct survey of enterprises (private institutions) asking them about their innovation activities, the Community Innovation Survey (CIS) being the best-known case. A good starting point for developing an original proposal about measuring innovation in the public sector would be that of setting the limits of a standard subject-based approach. In other words, one may wonder whether it is appropriate for the public sector to survey potential **innovators** (subjects), as it has been the case in the business sector; or it is rather preferable to replace or integrate such a focus on innovation subjects with a data collection scheme based on the measurement of **innovations** (objects) ⁽¹⁴⁾.

⁽¹⁴⁾ This discussion is partially echoing a similar debate that took place around 50 years ago between the measurement of business innovation by collecting information on the 'innovations' (new products or processes) and the development of dedicated business innovation surveys.

The high heterogeneity of public institutions (which should be the **subjects** of a survey about innovation in the public sector) has prevented, so far, the adoption as a standard of innovation surveys in the public sector. Even though public sector entities are asked to serve a common aim — public interest — in the most effective way, size, structure and tasks of public institutions are, quite often, hardly comparable within a single country and even less at an international level ⁽¹⁵⁾.

According to the European System of National and Regional Accounts ESA 2010 (p. 44), four groups of institutions are included in the ‘general government’ sector (ESA code S.13):

- i. central government (excluding social security funds) (S.1311);
- ii. state government (excluding social security funds) (S.1312);
- iii. local government (excluding social security funds) (S.1313);
- iv. social security funds (S.1314).

Of course, this breakdown is of poor use in statistical analysis by considering the many different typologies of institutions which belong to, at least, one of the three main groups listed above. This leads to the question about which sources can be used in order to identify the target population of a potential survey on public institutions and which official statistical classifications could apply to such population.

In terms of population of reference, it would be highly recommendable to use ESA 2010 as main reference, thus to adopt the definition of ‘general government’ ⁽¹⁶⁾. For both statistical and administrative purposes, EU countries are producing and regularly updating a list of institutions belonging to the ‘general government’ sector (ESA code S.13) as described in ESA 2010 (i.e. institutions and bodies supporting more than half of their activities with public funds).

As a conclusion, the ‘general government’ (S.13) population is probably the only comparable population of public institutions which can be used for consistent surveying activities across the EU ⁽¹⁷⁾. As a second step after the identification of several tens of thousands of bodies belonging to ‘general government’, a set of classifications have to be used to produce a relevant and meaningful output from any collection of data.

Table 2 provides some basic information about data available at national level. First of all, the size of the different sub-groups of the S.13 register can be very different across the EU as a consequence of the different structure of European states. France and UK, for instance, still have a much more centralised structure than Spain or Italy. As a consequence, specific public functions could be undertaken by a ‘central government’ institution in France and by a cluster of smaller local institutions in Italy. This, of course, strongly affects any comparability of production, and even more innovation, efforts carried out in the same area in the two countries. Second, the consistency of classifications will have to be checked. In principle, the national accounts’ classification by Functions of Government (COFOG) is the most useful for comparability needs but it is hardly made available by statistical agencies as it could be quite generic, with many institutions serving more than one function. On the other hand, also the classification by main economic activity (from French Nomenclature statistique des Activités économiques dans la Communauté Européenne NACE) is seldom used and could not properly describe the multiple tasks of a public institution. Moreover, a full consistency of the COFOG and NACE codes is hard to be achieved.

⁽¹⁵⁾ To be argued more extensively in the final report.

⁽¹⁶⁾ This concept is mentioned also in the Article 126 and Protocol 12 of the Treaty on the Functioning of the European Union (2008/C 115/01) in order to identify the institutions subject to the provisions given in the Treaty about the excessive deficit procedure.

⁽¹⁷⁾ The option of extending this population by including, for instance, the private providers of services of general interest cannot be excluded. Unfortunately, this will imply a detailed identification of such providers, based on national rules and market structures, which could affect the ultimate comparability across countries of the resulting lists.

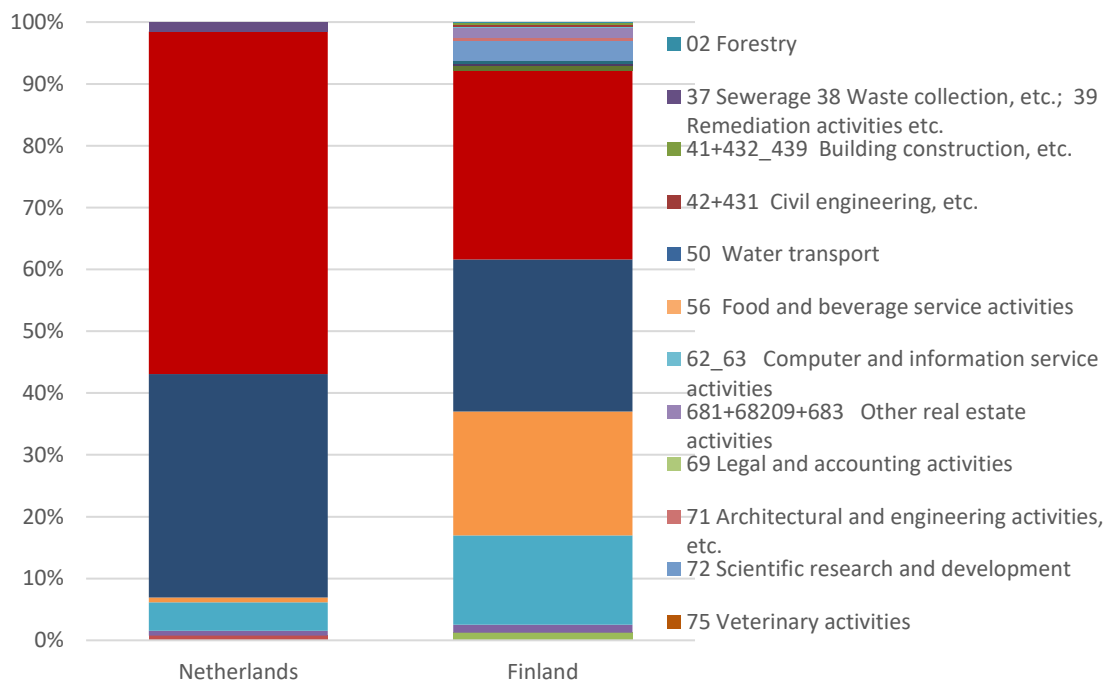
Table 2: Information on the publicly available lists S.1311 ('central government') in four selected EU countries, 2016.

| Countries | No of bodies | Managing institution | Nat. accounts | COFOG | NACE |
|-----------|--------------|--|---------------|-------|------|
| France | 735 | INSEE, National Statistical Office | No | Yes | No |
| Italy | 170 | ISTAT, National Statistical Office | No | No | No |
| Spain | 455 | Government of Spain, Ministry of Economy | Yes | No | Yes |
| UK | 869 | ONS, National Statistical Office | Yes | No | No |

Source: StarPIN data.

One of the key issues when targeting public institutions broken down by NACE is the different composition of the reference population across countries. The degree of specificity of national models can be inferred by considering the breakdown of gross value added calculated for the national accounts' purposes (Figure 6).

Figure 6: Breakdown of gross value added produced by the General Government sector by NACE. Year 2012. Finland and the Netherlands.



Source: Eurostat [nama_10_gdp](#)

As the provision of this information to Eurostat is voluntary, data are available only for a few EU countries but the comparison between Finland and the Netherlands could be sufficient to highlight to what extent tasks and structures of the public sector can differ between countries. In the Netherlands, General Government accounts for 14 % of GDP and is basically focusing on three main economic activities: public administration (including defence and social security) 55 %, education 36 % and residential care/social work 5 %. The share of government's value added in Finland is higher (21 % but is even higher in other EU countries) and is distributed across economic activities in a different way. Public administration is still relevant but covers only 30 % of total value added followed by education (24 %) and health (20 %, as it is managed by public institutions rather than by private institutions as in the Netherlands). Additionally, residential care/social work accounts for 14 % of the value added and a large public R & D sector is contributing, as well, with 3 % of value added. If the same comparison should have been replicated in terms of number and activity of public institutions, one can expect similar discrepancies, for instance many more public hospitals and public research centres in Finland than in the Netherlands. This is a clear evidence of the problems to be expected by comparing two mixed populations of public institutions for surveying purposes.

The same point can be reinforced by considering, for instance, the breakdown of government spending in EU countries by functions.

The COFOG approach, by considering activities to be undertaken by advanced countries more or less on a compulsory basis, allows for a higher homogeneity than the approach by economic activity (NACE).

Nevertheless, there are huge differences in the share of some specific functions between countries. Among others, Health — ranging from 6.4 % in Cyprus to 17.8 % in the UK, Defence — from 0.7 % in Luxembourg to 5 % in the UK, Education — 7.8 % in Greece and 16.2 % in Latvia, or Social protection (including pensions) — from 29.8 % in Cyprus to 44.9 % in Finland. Also for the COFOG approach, its translation in terms of number of institutions could generate huge differences on the composition of national samples or sub-samples of public institutions.

An additional point, which is often neglected in the discussion about the comparability of public institutions in different countries (political structures, cultural contexts, etc.), is the one referring to their differences in the organisational structure and the relative functioning practices. It could be assumed that the institutions belonging to areas where (international) mobility of personnel is easier, or even encouraged, are largely comparable, for instance across the EU. This is the case for public universities (even though their size also matters) or for public hospitals, as well as for a few other cases. On the other hand, a rationale expectation is that other institutions had been shaped in order to fill specific national needs and any comparison should be supported by a preliminary assessment. This is the case with primary and secondary education institutions, organisations providing housing or paying for social security instalments, centres for assistance and social work, etc.

As far as an international survey of public institutions should be planned, the following recommendations could thus apply:

1. Assuring, as much as possible, a consistency with concepts and definitions used in the System of National Accounts (main reference: ESA 2010).
2. Using official statistical classifications (for instance, by avoiding the national classifications of institutions by funding or overseeing ministry)
3. Limiting the comparison to institutions included in the ESA 2010 S.13 sector (General Government) or one or more of its four subsectors. Groups of business enterprises providing public services in different national contexts are hardly comparable.
4. Using 'functions of government' (COFOG) as a main classification scheme in order to improve comparability even between governments with different structures and priorities.
5. Avoiding using NACE codes for comparison between institutions, or groups of institutions in different countries. Public institutions' activities are often mixed and complex: as such poorly

described in NACE terms.

6. Taking into account national organisational and cultural specificities when comparing institutions operating in the same sector of the public administration but in different countries.

Which are the conditions for adopting an ‘object’ approach to measure public innovation?

So far, an extreme heterogeneity has been observed among institutions involved in public service innovations with the consequence of having to address several complex methodological issues when planning for an international survey of innovation practices and strategies implemented by public institutions (and even more when attempting to measure the actual outcome of their innovation projects). Given these analytical and methodological difficulties, an alternative choice could be taken into consideration: surveying ‘innovations’, rather than ‘innovators’.

As observed earlier, and stressed in a growing literature, the main goal of the public sector is to deliver high quality services to citizens (as well as to improve efficiency and effectiveness of service providers) — product and process innovations largely concern ‘services’. Thus, ‘public services’ could be seen as an alternative statistical unit to be targeted in order to identify changes and improvements as outcomes of innovation processes.

The first step will have to be, also in this case, a proper definition of the statistical unit. The definition ‘public services’ for services produced and delivered by a public administration is not commonly used in the EU legislation and documentation as alternatives concepts have been developed to catch specific phenomena.

A key concept emerged as a tool for the EU single market policy is that of ‘service of general interest’ (SGI) ⁽¹⁸⁾.

Services of general interest are services that public authorities of the EU Member States classify as being of general interest and, therefore, subject to specific public service obligations ⁽¹⁹⁾. They can be provided either by the state or by the private sector.

Examples of services of general interest include: public transport, postal services, and healthcare.

This group of services can be further broken down into three sub-groups.

Services of general economic interest (SGEI), which are basic services that are carried out in return for payment, such as postal services. These services are subject to European internal market and competition rules. However, there may be derogations to these rules if necessary to protect citizens’ access to basic services.

⁽¹⁸⁾ For reference, see the Protocol No 26 annexed to the Treaty on the Functioning of the European Union (TFEU): *PROTOCOL (No 26) ON SERVICES OF GENERAL INTEREST*

⁽¹⁹⁾ This absorbs the ‘... evolutionary concept of “universal service”, developed by the European institutions, [which] refers to a set of general interest requirements, which should be satisfied by the operators of such services to make sure that all citizens have access to certain essential services of high quality at prices they can afford [COM/2000/0580 and COM/2001/0598]. It is sensitive to national diversity and takes into consideration the special features of the European model of society.’ *Europedia* (http://www.europedia.moussis.eu/books/Book_2/3/6/06/4/?all=1)

Non-economic services (NES), such as the police, justice and statutory social security schemes, are not subject to specific European legislation or to internal market and competition rules.

Social services of general interest (SSGI) are those that respond to the needs of vulnerable citizens, and are based on the principles of solidarity and equal access. They can be both of an economic or non-economic nature. Examples include social security schemes, employment services and social housing.

This set of definitions allows for an improved characterisation of the services which are relevant for policy, in the EU arena. A key feature of them is that they do not have to be produced and delivered only by public institutions (i.e. general government institutions, as defined above). Thus, what allows for their identification is their relevance in terms of effects on the EU citizens and, as a consequence, their relevance for the EU policy: ‘services of general interest indicate “market” and “non-market” activities, considered to be of general interest by the public authorities, and subjected for this reason to specific public service obligations. Article 106 of the Treaty on the functioning of the EU (...) states that undertakings entrusted with the operation of (market) services of general economic interest are subject to the rules contained in the Treaties, in particular to the rules on competition, in so far as the application of such rules does not obstruct the performance of the particular tasks assigned to them (postal, telecommunications, transport, electricity, broadcasting, etc.). A Commission Decision sets out the conditions under which State aid in the form of public service compensation granted to certain undertakings entrusted with the operation of services of general interest.’⁽²⁰⁾ Both the European Union and EU countries are involved in the management of SGIs by considering their multifaceted role as outcomes, but also drivers of productivity, both of the public and the business sectors⁽²¹⁾.

This rich normative background left unaddressed the issue of defining SGIs in a clear and unambiguous way. The standard approach to fill this gap is that of using the standard economic classifications used in official statistics. As discussed above, the classification of government’s functions (COFOG) has to be kept as the main reference for international comparison purposes. On the other hand, the NACE classification — not always easily matched with COFOG’s divisions and groups — could be replaced by the Classification of Products by Activity (CPA), where ‘activities’ are just those included in the NACE classification, so a full consistency among the three classifications can be assured (Figure 7).

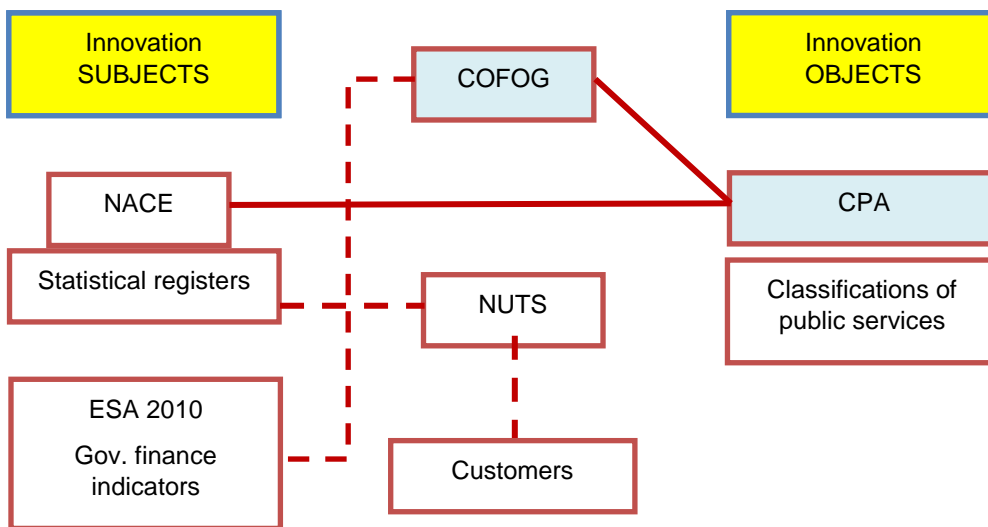
Figure 7 offers an original approach to the measurement of innovation in services as it is based on the assumption that an ‘object approach’ — i.e. focusing on services targeted by innovation activities — could be broader than only considering the potentially innovating institutions. In fact, a service can be identified by means of a CPA code which can be easily linked to both a COFOG and a NACE code allowing for the identification of the ‘innovator’. Of course, this is not true for the link from an institution to a service as an institution can be, by definition, a provider of multiple services and, as such, involved in process of innovating several services at the same time. The identification of the link between services (CPA) and government’s functions (COFOG), via NACE, is the methodological link of this approach. Unfortunately, while the CPA classification is fully consistent with the NACE classification, an official concordance table between COFOG and CPA does not exist⁽²²⁾.

⁽²⁰⁾ Eurostat, *ibid.*

⁽²¹⁾ ‘General interest services are a key element of the single market and of European competitiveness. They have to contend with significant changes in terms of the economy, technology and consumer requirements. To respond to these changes, Community action is designed to modernise these services to ensure that essential needs continue to be met and that performance is improved. Based on Article 106 (TFEU, ex Article 86 TEC), Community action takes account of the subsidiarity principle and tries to maintain a balance between market play and general interest in the form of universal service or public service obligations. The Commission has identified a number of guidelines in the field of general-interest services with the aim of formulating a coherent policy in this area [COM/2004/0374].’ Eurostat, *ibid.*

⁽²²⁾ Eurostat is currently working on the harmonisation of matrices used by several national statistical institutes in the EU matching CPA and COFOG codes for the national accounts needs. A methodological proposal is expected in 2018, so far national CPA/COFOG matrices are quite

Figure 7: Classification alternatives by adopting a subject or an object approach for surveying innovation in the public sector.



An official UN table of concordance between COFOG and economic activities has been produced with reference to the classification ISIC Rev.3.1 — roughly corresponding to the NACE Rev.1.1 — and, in order to be used as a bridge to the CPA 2008, it has to be translated into a COFOG — ISIC Rev.4 concordance table (where ISIC Rev.4 can be easily translated into the NACE Rev.2).

Before describing this process in detail, it could be helpful to identify those areas of service production in the public sector which have policy relevance.

In this perspective, the various areas of service production can be qualified and prioritised in terms of policy relevance (which is very important in terms of innovation policy). The priority areas are also those which should be preferably addressed in the measurement of innovation efforts, as they have a stronger impact on social and economic life in the EU.

A starting point could be the identification of **non-economic services** (NES), as they are services which are supposed to be made available to citizens in exchange for the contributions paid as general taxes and, usually, not requesting a further payment for receiving the relative benefits. In terms of service ‘division’ (2 digit CPA), they include: Remediation services and other waste management services; scientific research and development services; veterinary services for livestock; tourism promotion; public administration, security, law and defence services; compulsory social security services (the bulk of this group); and education services. The role of innovation in this area is mainly that of improving efficiency that is to give the community high level services at acceptable costs.

Another subset of services whose costs are expected to be covered by general taxation is that of the so-called **social services of general interest** (SSGI). According to the official definition, these services play a crucial role in improving quality of life and providing social protection for all EU citizens. They include: social security, employment and training services, social housing, child care, long-term care and social assistance services. A translation of this short list in statistical terms is anything but straightforward. In fact, social security is not univocally identified in statistical terms (neither in the CPA, nor in the COFOG classifications). In this respect, human health services cannot be totally included in the SSGIs but most of the social security services are included in the big

heterogeneous, starting from a 60 × 69 matrix in small countries to very large matrices at 3digit COFOG and 5digit CPA adopted in the largest EU countries.

division of 'Public administration and defence services; compulsory social security services'. Innovation is a key strategy to give access to these services to the most vulnerable part of the EU population (children, disabled, unemployed, elderly persons, etc.) and to assure their quality and effectiveness.

The largest group of services potentially produced by the public sector is that of the **services of general economic interest** (SGEI), largely encompassing also the SSGIs. In this case, the public sector is not expected to be the sole funder of the service provision but rather to make them accessible, at an affordable cost, to all citizens or, at least, to a large part of them. SGEIs include the collection, treatment and distribution of water, the distribution of gas, as well as the production and distribution of electricity, sewerage services and waste management, administration and, when needed, management of public transportation, public media broadcasting ⁽²³⁾, postal services ⁽²⁴⁾, communal services, education, health, general administration, cultural services. Innovation in these areas is largely involving business enterprises both as providers of new technologies and new equipment to public institutions and as key actors in developing new services.

A final group of services to be considered includes services offered by public institutions (or purchased from external providers) for internal use. This does not mean that they are not important for the functioning of the public sector and, even more, for allowing the public sector to make key services available to citizens. In this group, many technical services have been included — for instance, architectural and engineering services — which are needed to plan for, to build and to maintain buildings, infrastructures, and plants essential to assure a broad range of services to the EU communities. Innovating these services is, once again, an investment in higher efficiency and productivity.

There are, potentially, several hundreds of services produced, delivered or used by the public sector, which could be targeted with specific innovation strategies. In this respect, there are two implications to discuss. First, what is the most appropriate level for addressing the innovation issue and its measurement. Second, whether the innovation activities regarding a set of services should be prioritised, as both political objectives and phenomena to be measured. These are the conditions for developing a statistical tool for measuring innovation in the services produced or delivered by the public sector which could have the same quality of a survey on public institutions.

⁽²³⁾ This is an interesting example of the approach used by the EU institutions. While being apparently a non-essential need, the access to information through a public broadcasting service and media pluralism are seen as important conditions for a pluralist society and are directly related to the democratic, social and cultural needs of the EU citizens (as stated in the Protocol 29 annexed to the treaty of Lisbon).

⁽²⁴⁾ Postal services have been left to private business enterprises across the EU, nevertheless the EU legislation [Directive 97/67] defines a 'universal postal service' which must be guaranteed throughout the EU, while setting limits to the postal monopoly and providing a timetable for the controlled opening of the postal market to competition.

As far as a data collection on the innovation activities affecting **Services of General Interest** (and similar services) produced and/or delivered by public or private institutions is concerned, the following methodological recommendations apply:

1. Assuring, as much as possible, a consistency with concepts and definitions used in the System of National Accounts (main reference: ESA 2010).
2. Using 'functions of government' (COFOG) as a main classification scheme in order to improve comparability even between governments with different structures and priorities.
3. Considering both services produced and/or delivered by public institutions and services of general interest produced and/or delivered by private institutions or businesses.
4. Using official statistical classifications of products in the EU (CPA) as the main reference for the identification of services.
5. Testing the use of different data collection methods — surveys, collection of systematic information from administrative sources, collection of systematic information from the Web, collection of non-systematic information from the Web — in order to get the main advantage by the change in the statistical unit (from the service provider to the service itself).
6. Taking into account national organisational and cultural specificities when comparing similar services of general interest produced and delivered in different countries.

The services of general interest as statistical units: a CPA-based classification

A major achievement of this project has been the attempt to **match the COFOG and the CPA classifications** in a systematic way, thus filling a gap in this respect. This is an important step forward as it allows to define a standard which could be further improved and enriched with additional information provided both by statisticians (e.g. national accountants already working on this matching) and users (e.g. scholars, practitioners and policymakers involved in the regulation of the services' market in the EU).

For the purposes of this report, the standard COFOG classification has been taken into consideration at its highest level of detail: **classes (3 digits)**. The broader levels of COFOG are divisions (1 digit) and groups (2 digits). See Table 3.

Table 3: Overview of COFOG divisions and groups

| Government broad objective (division) | Sub-items |
|---|--|
| General public services | Executive and legislative organs, financial and fiscal affairs, external affairs; foreign economic aid; general services; basic research; R & D related to general public services; general public services n.e.c.; public debt transactions, transfers of a general character between different levels of government. |
| Defence | Military defence; civil defence; foreign military aid, R & D related to defence; defence n.e.c. |
| Public order and safety | Police services; fire-protection services; law courts; prisons; R & D related to public order and safety; public order and safety n.e.c. |
| Economic affairs | General economic, commercial and labour affairs; agriculture, forestry; fishing and hunting; fuel and energy; mining, manufacturing and construction; transport; communication; other industries, R & D related to economic affairs; economic affairs n.e.c. |
| Environmental protection | Waste management; water waste management; pollution abatement; protection of biodiversity and landscape; R & D related to environmental protection. |
| Housing and community amenities | Housing development; community development; water supply; street lighting; R & D related to housing and community amenities; housing and community amenities n.e.c. |
| Health | Medical products, appliances and equipment; outpatient services; hospital services; public health services; R & D related to health; health n.e.c. |
| Recreation, culture and religion | Recreational and sporting services; cultural services; broadcasting and publishing services; religious and other community services, R & D related to recreation, culture and religion; recreation; culture and religion n.e.c. |
| Education | Pre-primary, primary, secondary and tertiary education, post-secondary non-tertiary education, education non definable by level, subsidiary services to education, R & D; n.e.c. |
| Social protection | Sickness and disability; old age; survivors; family and children; unemployment; housing; R & D; social protection and social exclusion n.e.c. |

The matching process has started from the COFOG structure and from the identification of the economic activities linked to each COFOG class; as already mentioned, the official UN Table of correspondence for the COFOG and ISIC Rev. 3.1 classification has been used, taking into consideration **135 COFOG classes** and matching them with the most compatible ISIC Rev. 3.1 code, usually at 4 digit level (classes) but also, when more appropriate, at 3 digit (group) or at 2 digit (division) level.

As a second step, the ISIC Rev. 3.1 classes have been translated into ISIC Rev. 4 classes. ISIC Rev. 4 has been published in 2008 in order to replace the Rev. 3.1 dating back to the year 2002. As the link between the codes in the two classifications is not of a one-to-one kind, the 135 ISIC Rev. 3.1 codes have generated **223 ISIC Rev. 4 codes**, most of them at 4 digit (classes) level, as well.

In order to allow for a full consistency with the EU classification standards, the ISIC Rev. 4 codes

have been further translated into NACE Rev. 2 classes, which is the version of the statistical classification of economic activities developed for the EU needs. The substantial consistency between the two international classifications made the process quite straightforward.

The final step has been that of checking for the logical consistency between COFOG and NACE codes and, at the same time, to identify the CPA codes matching at best the government functions described by the COFOG codes. Of course, not all the 'products' (in terms of CPA codes) included in a single NACE class have been taken into consideration if not significant for the public sector, i.e. not representing services of general interest.

A key choice has been that of using the most disaggregated definition of CPA 2008 products (services, in this context), i.e. the sixth level (6 digit).

CPA has a hierarchical structure with six levels, each identified with a specific code:

First level: 21 sections (alphabetical code);

Second level: 88 divisions (two-digit numerical code);

Third level: 261 groups (three-digit numerical code);

Fourth level: 575 classes (four-digit numerical code);

Fifth level: 1 342 categories (five-digit numerical code);

Sixth level: 3 142 subcategories (six-digit numerical code).

Out of 3 142 products described in the CPA, **412 subcategories** (almost all of them services) have been identified as products of general interest and linked to a function of government at class level. The matching has resulted in a matrix of **135 COFOG classes** and **412 CPA subcategories**. The relationship can be also described as a non-biunivocal table of correspondence (see Annex 1).

The key methodological point to be highlighted here is the level of detail available in CPA for the description of services of general interest. On the one hand, it is not possible to recommend an official statistical definition more detailed than that used in this matching. On the other hand, the relevance of the information provided in Annex 1, both with reference to the description of products/services and the consistency with COFOG classes will have to be thoroughly tested on the field when evidence of service provision and service innovation will be available.

6

Methodology for data collection and data analysis

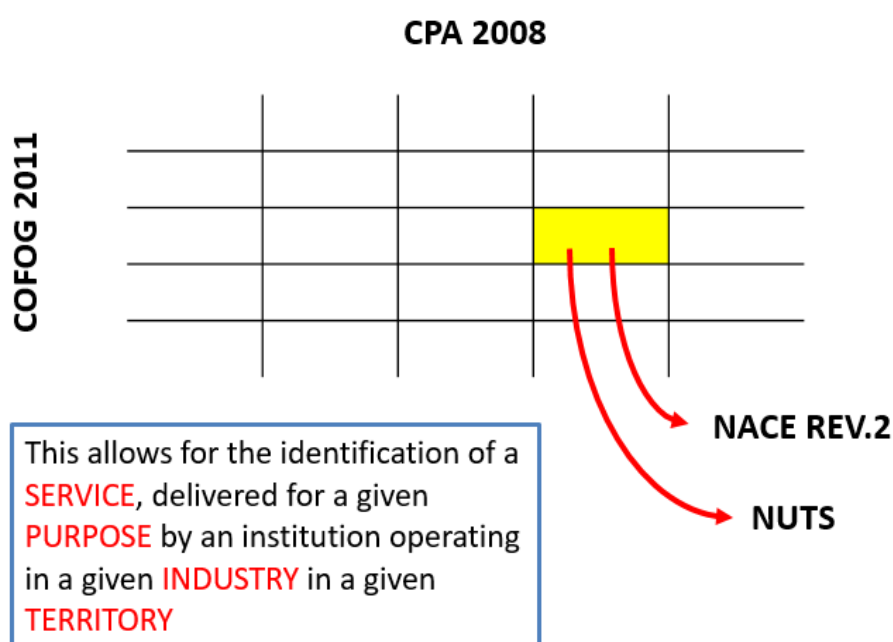
This chapter illustrates the process through which services are identified and characterised in terms of innovativeness, and through which data are collected and organised. While the analysis was conducted in view of its application to specific public sector domains, the method described can be applied to any category of public services. The steps described include: the identification of public services, in correspondence with extant classifications, and of institutions involved in their provision; the description of service characteristics that can be associated with different levels of innovativeness; the singling out of descriptive tools (see Annex 3 for a list of keywords and of logically grouped sets/‘baskets’ of keywords) that can be managed via web-scraping techniques; the running of web-scraping using appropriate techniques; the organisation of results of data obtained through the web-scraping process; the validation process through the use of complementary data sources, including administrative data. A description of the web-scraping source is provided in §6.2 and details of the preparation activities for setting and running the tool are organised in the User manual (in Annex 2).

The process

The phases of the process for collecting data on public service innovation with the additional source provided by the web are the following.

A. Identifying and selecting the public service

On the basis of the classifications of public functions, public institutions and public services, the service to be analysed should be selected and would need to correspond to one or more of the cells of the classification matrix.

Figure 8: The classification grid**B. Identifying institutions involved**

Thanks to experts' knowledge and accessible studies of the service evolution, the institutional organisation/s backing the selected service are identified. A list of websites is retrieved for the relevant institutions to explore service characteristics and the extent of their provision.

C. Capturing the degree of service innovativeness

The levels of service innovativeness are defined in terms of the qualitative characteristics of the service. Information on the basis of actual and perspective characteristics of services are obtained from sectoral studies, analyses of the technology and organisational evolution of the areas concerning the selected service, and studies of policy goals affecting the search for service quality and performance improvement. Qualitative characteristics of services are associated to different levels of innovativeness (different 'rungs' of the 'ladder') at a given moment in time, and changes in characteristics are associated to changing levels of innovativeness over time (movements from one rung to another along the ladder). See Chapter 1 for a description of the 'ladder approach'.

D. Defining the tools for web-scraping

Once the list of URLs of the relevant institutions/public bodies responsible for providing the service is built (see point B above), and the areas and directions of innovativeness in the specific public service considered are traced (see point C above), a set of keywords is selected with the aim of capturing the innovativeness levels of services on the website of the institutions providing it. The presence of the keywords should allow to answer the question: 'which innovativeness degree can be assigned to the service X provided by institution Y?'. Keywords can be either generated *ex ante* by experts and researchers based on prior knowledge or obtained via automatic search techniques applied to the exploration of websites of the institutions involved in service delivery. At any rate, some involvement of human agents is needed to select and enrich keywords either in the *ex ante* or *ex post* phases, or in both. The StarPIN methodology has consisted in preselecting a set of keywords based on expert advice, and enriching them through a testing phase of search in relevant websites.

More precisely, the keywords are identified and enriched accessing to a pre-determined subset of virtuous and of lagging behind institution websites using a manual procedure assisted with an electronic search procedure (using tailor-made scripts based on the open source Node environment).

As many as possible synonyms of keywords and different writing options are added to the list. The keywords are organised in a matrix with corresponding translation in English and in each of the pilot country languages. The two starting languages in the present application experiment are Italian and Dutch, mediated by English, then followed then by Portuguese, Slovak and Swedish (see Chapter 3 and Annexes 3 and 4 for an illustration of the use of keywords to detect innovativeness in selected service domains).

The innovativeness levels were initially identified in correspondence with the appearance of individual keywords in the website. Single keywords often led to an excess of false positives for high levels of innovativeness. To reduce these false positives, suitable combinations of two or more keywords have been identified.

In the first pilot service examined — waste collection provided by municipalities — an example of keyword that was raising some issues in terms of false positives was represented by the keyword ‘identification code’, which was considered a signal that waste management was conducted with a high degree of waste separation (i.e. tied to an individual user), hence a relatively high level of innovativeness. However, ‘identification code’ also refers other issues, like public procurement. Combining ‘identification code’ with ‘paper waste’ allowed for better identifying a given level of innovativeness in waste management. This is the main rationale for introducing the ‘basket’ approach.

Each keyword is attributed to one and only one basket. The presence of one of the words of one basket reflects the fact that the aspect described by the basket is mentioned in the analysed page. Baskets are ‘turned on’ through the observation of at least one keyword associated with the corresponding basket code. This concept reflects the Boolean operator ‘OR’: if any of the keywords that belong to the basket is found in the webpage, that is enough to confirm that the aspect described by the basket is attributable to the specific service (see Annex 3 for an illustration of the organisation of keywords into baskets in the two service domains under observation).

The presence in a given webpage of at least one keyword from all the different baskets that describe a specific level of innovativeness identifies the ‘AND’ Boolean operator: both basket n AND basket m are present. Different combinations of baskets identify different innovation levels.

What follows is the short description of the aspects to be ‘captured’ by the subgroups of keywords organised in baskets for the pilot service waste collection:

- i. Waste, waste bins, waste collection (baseline characteristic: identifies that a waste management service exists);
- ii. ‘Structural’ variables which influence the tariff (house size, no of inhabitants);
- iii. Debris or bulky waste (not strictly related with the service evolution, but relevant in some countries, namely NL);
- iv. Home collection based on waste ‘fetching’;
- v. Separate collection;
- vi. Waste collection based on waste ‘bringing’;
- vii. Home collection based on waste ‘fetching’ with personalised systems;
- viii. Measurement of waste amount based on the number of bags/volume;
- ix. Identification system based on information encoded in a chip;
- x. Personalised systems of waste ‘bringing’
- xi. Measurement of waste amount based on weight/weighing systems;
- xii. Pay-as-you-throw, including tariff systems (variable tariff/tariff rate).

To capture the ‘AND’ operator, the coexistence of baskets which are found on the webpage is checked for in each of the pre-selected URLs. A basic combination was used to detect the level of innovativeness of a service and to attribute a level indicator to the institution providing that service. In

our first pilot case on waste collection, the first set of combinations tentatively assigning the innovativeness level is:

1 (2 adds details): the URL contains information related to the description of waste collection, with no further details on service characteristics = **innovation level 1**

4 and 5: evidence of separate waste collection and of garbage 'fetch' at home = **innovation level 2.a**

5 and 6: evidence of separate waste collection and of garbage 'bring' practices = **innovation level 2.b**

1 and 7, 1 and 8, 1 and 9: the thrower is identified with a chip, bags/volume counted, with waste collected at home ('fetch') = **innovation level 3.a**

1 and 10, 1 and 11: the thrower is identified, bags/volume/weight counted, with garbage 'bring' practices = **innovation level 3.b**

1 and 12: evidence of pay-as-you-throw practices = **innovation level 4.**

It may be debatable that '2a' level identifies a higher innovativeness than '2b', and the same applies to '3a' vs. '3b'. The degree of innovativeness may depend on the value attributed to different options which could also be context-specific. At this stage, the four levels (2.a, 2.b, 3.a and 3.b) can be kept as such, leaving to data provider the choice whether using these distinctions and, possibly, ranking them differently according to their specific needs. In some member countries the 'a' vs. 'b' juxtaposition may be perceived differently, and identify higher or lower value creation in terms of user friendliness of the service provision and/or externalities generated. In other circumstances, service providers as well as users, actors and institutions involved may be scarcely affected (if any) by this characteristic of the service. A simplifying choice might well be to stick to the four 'main' levels of innovativeness, namely 1, 2, 3 and 4, allowing for the possibility of further unbundling levels 2 and 3 to account for the 'fetch' or 'bring' characteristics, whenever this makes sense or has a relevance for analytical or normative reasons. In some of the tables below both the detailed and the simplified ranking will be used.

For what concerns software, languages, scripts, machine and connection, an ad hoc 'library' of instructions for an automatic searcher on the web was generated by the StarPIN project team, and is described in the User manual (Annex 1).

E. Running the web-scraping

The keyword search for service X in the URLs of institutions Y is launched using automated techniques (see §2.2).

F. Organising the results according to the innovativeness indicator and analysing the process outcomes

A spreadsheet was created for organising the web-scraping results, by grouping the data collected by webpage by domain/institution, and by identifying the level of innovativeness is detected through the presence of at least one keyword from each of the baskets identified according to the methodology described above (point D). So, once the search is completed, the results are stored in the spreadsheet file, to obtain an indicator of innovativeness for the service provided by each institutional subject/actor.

The 'result' of the web-scraping is the attribution of a level of development of the selected service, through innovativeness steps. The development level reached can be considered reliable and powerful as it captures a substantial part of the innovative phenomenon: the characteristics of the service provided, captures through the words used to describe and provide and manage the service.

The process described is consistent with, and allows for, additional specifications and introduction of further 'new' levels of innovativeness as long as they emerge and can be described/distinguished from the previously identified ones. The method is versatile because the content (keywords) is strictly split from the process (the web-scraping technology being used and the processing of the data).

This method of data collection and management might be used in sequence and/or in parallel with other methods based on censuses/samples/surveys. At this stage, the indicator of innovativeness level can be used together with administrative and statistical data to validate it and to perform cross-sectional analysis. The indicator created should be inserted in a table together with data for validating the web-scraping results, and data for cross sectional analysis selected on the basis of their relevance to the innovation phenomenon and ordered by the service provider/public subject/entity and geographical/administrative unit.

Validation phase and cross-sectional data analysis

The innovativeness level score obtained through the described process is then ‘controlled’ checking its coherence with administrative and statistical data related to the service studied and with the actions and phenomena related to it, statistically detected for correspondent geographical and institutional levels.

Current status of site-centric measurements

In this section, a set of viable options to collect, process, classify and evaluate information from websites from public sector organisations is described. This description is based on the first results of a pilot measurement on two particular public services (waste collection and Health homecare) in two countries (Italy and the Netherlands). The procedure has subsequently been extended to three additional EU countries, namely Portugal, Slovakia and Sweden. Some illustrative results of this measurement exercise for these countries are discussed in Chapter 7.

Data collection

Overview of web collection methods

Broadly speaking, there are two basic methods to collect data from the internet. They differ in the place where the intelligence in the implementation resides. One can either (a) put the intelligence in the script that collects data from websites or (b) download the entire content of data (‘site mirroring’) and then filter the data afterwards (Brennenraedts et al., 2012). The advantage of the first method is that the classification of data is already (partly) being done during the process of data collection. As a result, only a selected part of the data needs to be downloaded and stored. The disadvantage is that the selection is tailor-made to the particular aim of the measurement. Hence, if the selection is too restrictive, or if the aim of the measurement shifts, the entire data collection process has to be re-done.

Which method (or technique) — or combination of techniques — works best is very much dependent on the particular characteristics of the units of observation, such as the number of websites involved, the size of websites (number of web pages), the structure of the website (semi-structured or ill-structured) and the type of information that needs to be identified. Because this can only be evaluated case by case, different techniques have been tested, in parallel, in the pilot study: two instances of the first method (the use of search engine results, and web-scraping) and one instance of the second method (website mirroring).

Pros and cons of the methods

An obvious advantage of site mirroring is that all data is available offline. Therefore, any type of selection or cross-section of the data can still be made afterwards. The download of entire websites often involves massive amounts of data, hence scalability is a critical issue. With an average size of several hundreds of megabytes this will soon turn into large datasets. This severely increases overall download times (days, up to weeks), requires massive storage capacity and fast computers to sort through the amount of data. The storage, processing and filtering of such large amounts of data requires a dedicated IT-infrastructure (either locally or in the cloud), and the skills to use such infrastructure. Searching through a large dataset with a standard computer and standard consumer solutions takes a long time and is not a viable solution. To do text-based search on data of this magnitude a server (or cluster of servers) is needed with a specialised software application. A

commercial specialised service provider, namely ElasticSearch, has been used for this task ⁽²⁵⁾.

In order to make data handling possible while only using a basic IT-infrastructure, the amount of data that would be copied has been restricted to a given volume. A few filters, more or less similar to those used for web crawlers (see hereafter), have been used for this purpose:

- Limit data collection to the first three levels of the website (some public sector organisations host a lot of historic data and to download and process the data would take a lot of time and hence disrupt the workflow. Three levels seem to give the best trade-off between size and loss of potentially relevant information).
- Only download HTML content (allow only files that have text: html, php, txt.). The assumption is that services should already be sufficiently described on the website itself (hence not in separate files, e.g. word or pdf files).
- Ignore all files that are > 10 MB.

For the selective mirroring of website, HTTrack (<https://www.httrack.com>) has been used. It also works for larger sets because the process runs fully autonomously. HTTrack automatically stores websites locally in such a way that the results are browsable like a normal (online) website. The program 'redirects' all links to the downloaded content. A major drawback of HTTrack is that the results occurred not to be fully reproducible (i.e. subsequent runs produced different results for the same website).

For the Netherlands alone, about 100 500 files have been downloaded with the .html extension. However, throughput times were still substantial — up to several hours for one website. Moreover, about 5 % of the files could not (or only partially) be downloaded.

A possible variant of the first method, that is putting the intelligence in the script that harvests data from websites, is the use of **search engine results**. This method has the advantage that all processing is being done by the search engines, and that a lot of additional — and potentially relevant — metadata is being added. The key advantage is that results not only return the targeted webpages for a specific public sector for a specific public sector organisation but also webpages from external organisations that are conceptually also related to the same service. This enables the positioning of the public service organisation in a broader context (e.g. whether there is a lot of public–private partnership in the provisioning of the particular public service, or whether several public sector organisations use shared services, i.e. have the same partner).

The use of search engine results has however become so popular that most owners of the search engines have legally and technically restricted the use of automated scripts to acquire results. This is a major drawback. It basically rules out the use of leading search engines such as Google and Bing and forced the team to use less popular search engines such as DuckDuckGo. The coverage of these search engines is lower than Google and Bing. However, the difference is likely to emerge only in small websites (that are not yet being indexed by DuckDuckGo) and most government websites will still be included. Nevertheless, the automated use of DuckDuckGo also soon ran into limitations. These limitations had as a consequence to allow for retrieving only a modest amount of data. After a seemingly random amount of pageloads, it was not possible to connect to DuckDuckGo for a few days. Since the amount of requests was far below what was needed to achieve the intended results — and it was unknown whether the block of DuckDuckGo should be expected to be a permanent one — a decision had to be taken to give up with the search engine option.

⁽²⁵⁾ <https://www.elastic.co>

After this change of strategy, a further 15 000 links (which amounts to 750 page-loads) were collected. This was enough to cover the Dutch waste collection case, as the set is limited to the websites from about 400 municipalities. However, for the Italian case (with about 8 000 municipalities — with 10 keywords already adding up to about 80 000 page-loads) the limits set by DuckDuckGo were far too restrictive.

In short, the use of search engine results is an efficient manner to quickly test the validity of a small set of websites (which could be a random sample from a larger set). It also gives potentially relevant information about the context of public services, which is unique for this particular method. However, the scalability of the method remains a critical issue and needs to be further assessed.

Eventually, the second variant of the first method (**web-scraping**) was adopted. This is because the tailor-made scripts that were developed for the StarPIN project can be run on a standard computer by users without prior knowledge of big data processing techniques. In addition, they do not run into the limitations that are set by the search engines.

Application of the site-centric measurement in two countries

This section provides a close-up on the application of the methodology to the pilot public service waste collection of municipal solid waste (MSW), including web-scraping and the validation process.

Pilot service waste collection

A first run of site-centric measurement was carried out to identify the innovativeness of waste collection services in the case of Dutch and Italian municipalities. The samples used for this measurement exercise in the various phases of the experimental application of the web-scraping tool are described below, and some descriptive statistics are provided.

Table 4: Samples used for the measurement exercises

| | Italy | Netherlands |
|-----------------------------|-------|-------------|
| Nr of internet sites | 261 | 328 |
| Sites without hits | 26 | 2 |
| Average | 44 | 17 |
| Median | 23 | 16 |
| Maximum | 361 | 126 |
| Minimum | 1 | 1 |

Source: StarPIN data.

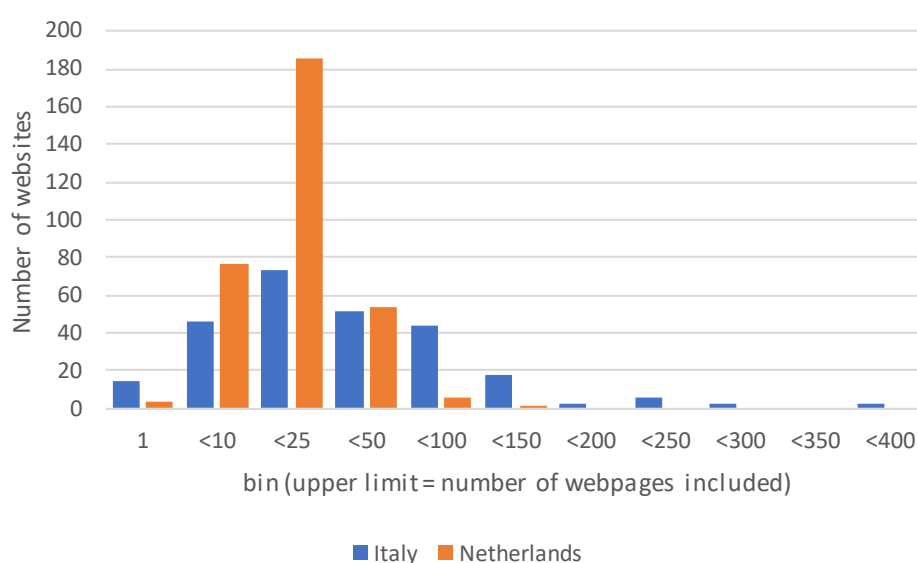
About one third of the Italian sample consisted of small municipalities. The size of these municipalities (in terms of inhabitants) is much smaller than the size of the smallest Dutch municipalities. This might largely explain the much higher number of Italian websites where no keywords were found at all. The average size of websites (in terms of number of selected webpages) is higher for the Italian than the Dutch set. This is mainly due to the presence of a small number of large websites (i.e. the Italian set has a 'longer tail' than the Dutch set has).

Table 5: Descriptive statistics for the Dutch and Italian set of municipalities included in the waste collection pilot

| | Italy | Netherlands |
|-------------------------------------|-------|-------------|
| Sample size | 261 | 328 |
| Total websites with keywords | 235 | 326 |
| Average | 44 | 17 |
| Median | 23 | 16 |
| Maximum | 361 | 126 |
| Minimum | 1 | 1 |

Source: StarPIN data.

Figure 9: Distribution of number of included webpages per municipal website, waste collection pilot, Italy and the Netherlands



Source: StarPIN data.

These preliminary findings led to the choice of focusing the Italian web-scraping to all the cities/towns which are capital of the corresponding province, and to the additional group of towns with more than 50 000 inhabitants, leading to a coverage of Italian inhabitants exposed to the waste collection public service analysed that is slightly more than one third of the total population.

Some preliminary results of the application of our measurement procedure for the Dutch and Italian sets (first run, n = 261) are shown below. Innovation level = 0 refers to the websites where no single keyword from the list was found. This is most likely due to technical issues (e.g. ill-functioning websites).

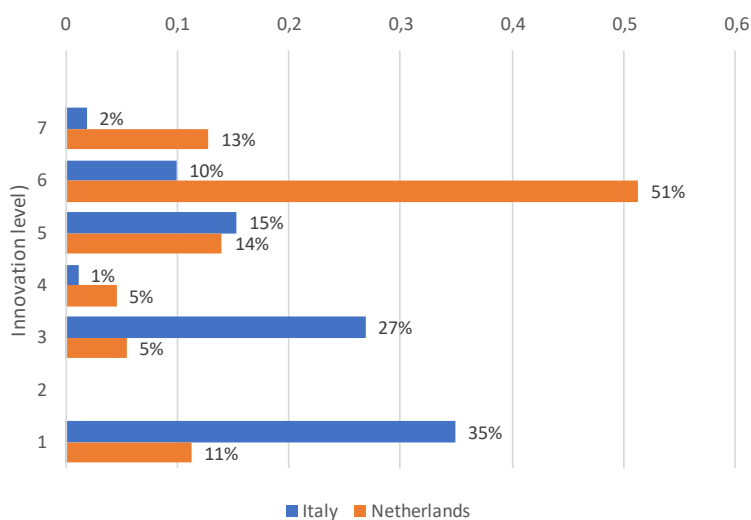
Table 6: Results for Italy and the Netherlands

| Innovation level | Italy | | Netherlands | |
|------------------|--------------|--------------|--------------|--------------|
| | % | <i>n</i> | % | <i>N</i> |
| 4 | 2 % | (5) | 13 % | (42) |
| 3b | 10 % | (26) | 51 % | (168) |
| 3a | 15 % | (40) | 14 % | (46) |
| 2b | 1 % | (3) | 5 % | (15) |
| 2a | 27 % | (70) | 5 % | (18) |
| 1 | 35 % | (91) | 11 % | (37) |
| *0 | 10 % | (26) | 1 % | (2) |
| Totals | 100 % | (261) | 100 % | (328) |

Source: StarPIN data.

For the current sets of municipalities, (whereas the Dutch sample covers nearly the entire population frame), overall the innovation level is much higher for the Dutch than for the Italian municipalities. The weighted average for the Dutch municipalities (5.1) is almost twice as high as the average for the Italian municipalities (2.7).

Figure 10: Scores on innovation levels for municipal websites, waste collection pilot, Italy (first run) and the Netherlands



Source: StarPIN data.

Note: rankings from 1 to 7 correspond to the more articulated scores used in Table 6 and elsewhere in this report (0, 1, 2a, 2b, 3a, 3b, 4), wherein increasing levels of innovativeness are assumed for 2a, 2b, 3a, and 3b. See section 2 above.

To assess their **(external) validity**, the results of a run are compared against a known set of positives (descriptions of a public service on a specific web page that are known to be on a specific innovation level) and a set of negatives (descriptions of a public service on a specific web page that

are known to **not** be on a specific innovation level) ⁽²⁶⁾. Moreover, information about the level of innovation can also be compared with actual measurement of performance at a cross-sectional level. For the Dutch waste collection case, information was collected about the following variables:

- i. Method of domestic waste collection by municipality (2015);
- ii. Differentiated tariff (Diftar) system by municipality (2016);
- iii. Total collected amount of fine residual waste (2015, Δ 2010-2015);
- iv. Total separated amount of plastics, cans, drinks carton (2015, Δ 2010-2015);
- v. Percentage of separated total household waste (2015, Δ 2010-2015).

The challenge is that the data quality of the benchmark data itself could also be improved. The first two (input) variables are closely related to the classification of innovation levels that has been used in the site centric measurements. The results of the tests should therefore correspond with these variables. Alas, the benchmark data is relatively outdated (2015 and 2016, respectively) and the current situation — as measured in the pilot (2018) — might in the meantime have changed considerably.

The last three variables are output variables with presumably high internal validity. The assumption is that municipalities that use the most advanced collection systems (hence have the highest innovation levels) perform best on the output variables (hence have the lowest amount of waste per inhabitant and the highest shares of separated collection). However, in this case, also, the data is relatively outdated.

This is a structural aspect of statistical data which requires time to be built, collected, aggregated, completed, and will always present a time delay compared to the results collected through web-scraping.

Keyword identification strategy

The identification strategy was carried out through a basic use of the Node tool ⁽²⁷⁾, in which there was no repeated execution of scraping for all municipalities. As a result, a small subset of municipalities was identified for which was scraped a wide range of keywords chosen from literature and from documents related to the waste collection. In the specific case of Italy, extant literature ⁽²⁸⁾ suggested the following keywords, also validated from a 'manual scraping' ⁽²⁹⁾ in each of municipalities analysed at this stage ⁽³⁰⁾.

Focusing on two dozen Italian municipality websites, a manual data search has been conducted to refine and enrich the keyword list based on the test of the results obtained via web scraping. The list of municipalities is reported in the following table:

⁽²⁶⁾ The **accuracy** of the results can be calculated as the ratio of the correctly predicted observations (the sum of all true positives and true negatives) to the total observations. Accuracy would be sufficient as a stand-alone measure to evaluate the quality of the results assuming that the underlying dataset is symmetric (i.e. values of false positive and false negative are quite similar). However, for the models used in this study this is most likely not the case. Therefore, two additional measures to evaluate performance are needed, namely **precision** (the ratio of correctly predicted positive observations to the total predicted positive observations) and **recall** (the ratio of correctly predicted positive observations to the all observations in actual class). Precision and recall can be combined into the so-called **F1-score**, which is simply the weighted average of the two measures. The F1 score is a good alternative to accuracy, especially when the underlying class distribution is rather uneven.

⁽²⁷⁾ NODE is an open source, completely free, JavaScript-based framework/platform to develop a variety of web applications.

⁽²⁸⁾ ISPRA report 2017 and the Italian EIONET-EEA report 2017.

⁽²⁹⁾ A simple manual desk research in each municipality website.

⁽³⁰⁾ At this level of analysis, the subset was chosen without any criteria considering the geographical macro-area, the population, etc. The choice can be considered purely random.

Table 7: List of manually searched municipalities

| | | |
|--------------------|----------------------|----------------|
| MULAZZO | PALERMO | GAGLIANICO |
| PRATO | NOVARA | MONTE CREMASCO |
| MONTEMIGNAIO | MONTECALVO IN FOGLIA | FERRARA |
| SARACINESCO | AGORDO | LONATE CEPPINO |
| LONGOBUCCO | BARGAGLI | VENEZIA |
| SALAPARUTA | FROSINONE | CAPRACOTTA |
| BALESTRATE | MILO | PORDENONE |
| CORIGLIANO CALABRO | BISCEGLIE | MATERA |
| TOMBOLO | SEGGIANO | SAN FELE |
| VIGHIZZOLO D'ESTE | ROVIGO | VITERBO |
| CALTIGNAGA | SAGRADO | BORGO PACE |
| MARANZANA | REGGIO CALABRIA | |

Source: StarPIN data.

After the identification process of the keywords, each of them was assigned to a basket as described above. See Annex 3 for a list of keywords and of baskets used to explore the innovativeness levels in the two pilot services.

Web-scraping process

The definition of keywords and baskets allows for a concrete web-scraping analysis using the NODE and PUPPETEER software in a Linux setting. The Italian local administrative division counts 7 954 municipalities; out of this population a sample of 150 municipalities was drawn. Preference was given to municipalities at province level and to those with more than 50 000 inhabitants. The selection allows for covering about 20 million inhabitants.

The web-scraping analysis started by considering the list of municipalities and keywords selected in a file and then running the software. For this aim, a library — a set of scripts which compose the steps for the robot search, and enables the programme to run — was created.

The detailed results are stored by the software in a .csv spreadsheet that contains every municipality 'subdomain' URLs with the count of the keywords found in each of the webpages. In addition, an aggregation of the results found in a subdomain setting — which is added up at a 'domain' website level — was performed. The further steps followed the basket approach presented above.

In some cases, the software failed the crawl of the municipality website either because the navigation timeout exceeded 80 000 ms or due to the impossibility of visiting the specific URL to be analysed.

The innovativeness indicator generated is then analysed and compared with administrative data related to the corresponding institutional and geographical 'dimensions'. In the case of Italy this has been done by comparing web-scraping data with administrative data published by municipalities on their environmental practices (as collected in the so-called MUD 'Modello unico di dichiarazione ambientale' (Unified environment statement form) page 56 and page 113 'G. U. n. 303 del 30/12/2017 variable "tariffa puntuale" (pay-as-you-throw) and with the available statistical data related to MSW at country or regional level. See below for details on this validations strategy.

Validation strategy

Using available administrative and statistical data, the results extracted through the web-scraping analysis underwent a process of validation. In particular, a check was undertaken on whether the innovativeness level data obtained through web-scraping were coherent with the data from MUD, a specific source of administrative data for the Italian case.

MUD represents an annual declaration presented by those who professionally carry out waste collection, transport activities and waste recovery and disposal operations. In this setting, also municipalities are obliged to submit this statement about ⁽³¹⁾:

- i. the quantity of municipal waste collected on their territory;
- ii. the quantity of special waste collected in their territory following a specific agreement with public or private parties;
- iii. the subjects responsible for the management of the waste, specifying the operations carried out, the types and the quantity of waste managed by each of them;
- iv. the costs of the technical and financial management and depreciation of investments in waste management activities, as well as income from the tariff referred to in Article 238 and income from consortia for the recovery of waste;
- v. data on separate collection;
- vi. the quantities collected, broken down by materials, in implementation of agreements with the consortia for the recovery of waste.

MUD summarises data about municipal costs and revenues related to the waste management. Hence if MUD indicates the availability of data about a variable waste collection tax or any other revenue related to the weight or volume of waste collected, then it could be inferred that there is a higher level of service innovation. In a similar vein, the cost of a service can reflect its level of innovativeness to the extent that it is associated with the provision of higher quality services offered by the municipality.

Summarising and integrating collected data

Table 8 illustrates the type of data available for waste management in Europe, including indicators collected from statistical and administrative sources. The table summarises statistical data from different sources about MSW that is possible to link to the level of innovativeness of our service. Using country and regional levels of analysis, it is possible to match the information across countries and regions in order to analyse differences and similarities. In particular, the share of sorted waste on total volume of waste could give an indication of the efficiency level of the service offered (i.e. the level of innovation to which it belongs). Data are collected from the main official sources: Eurostat, OECD, EEA and ISTAT ⁽³²⁾.

For example, the Eurostat database provides information about MSW from households and small enterprises at NUTS 2 level. In particular the dataset contains the share of landfilled waste which could be important information: if this share is high, maybe not a large amount of waste collected is differentiated.

Table 9 exemplifies how data collected through web-scraping on waste collection services in the case of municipalities can be elaborated to generate an indicator that aggregates the innovativeness level service on a different geographical scale (regional, or macro-area). This upscaling can lead to a detailed and comparable analysis with the existing statistical and survey data-source. (see Chapter 7 for an upscaling exercise which uses the analytical procedure illustrated in Annex 5). It suffices here to highlight that the logical structure used to exemplify how our innovation indicator can be combined with data from different sources at various levels of integration for the Italian case, can be replicated for other countries, thus filling in the corresponding lines in Table 9.

⁽³¹⁾ This is the general difference between 'MUD Comuni' (MUD Municipalities) and 'MUD produttori e gestori' (MUD producers and providers).

⁽³²⁾ The National Institute of Statistics — Italy.

Table 8: Description of data available on and related to waste collection from different sources, for EU, IT and NL

| Description | Unit of measure | Area geo | Statistical unit | Connections with public innovation | Link |
|--|-------------------------------------|---|---|---|---|
| Data on the generation and treatment of waste collected from the Member States. The waste management indicator set aims at showing how much of a country's or of the EU's own waste (in the following referred to as national waste) excluding major mineral waste is actually recycled, incinerated (with energy recovery and without), landfilled or backfilled. | Amount of waste in tons, percentage | EU Member State | The indicator covers non-hazardous (nh) waste from all economic sectors and from households | Data give information about what happens after waste is collected (after the service analysed) to recovery or to disposal (and in which way) at country level and not only for MSW | http://appsso.eurostat.ec.europa.eu/nui/submittViewTableAction.do |
| Data on the generation and treatment of waste is collected from the Member States. The information on waste generation has a breakdown in sources (19 business activities according to the NACE classification and household activities) and in waste categories (according to the European Waste Classification for statistical purposes). The information on waste treatment is broken down to five treatment types (recovery, incineration with energy recovery, other incineration, disposal on land, and land treatment) and in waste categories. | Tons of waste and in kg per capita | EU Member State | Non hazardous | In particular the presence of data about the most important categories of waste collection (different waste fractions: glass, plastic, rubber, etc.) gives information about the amount of different waste collected by households and different economic sectors even if data refer to a wider set of waste categories and not only to MSW | http://appsso.eurostat.ec.europa.eu/nui/submittViewTableAction.do |
| Municipal waste: Waste from households and small enterprises | 1 000 tons | All municipal waste in a NUTS 2 region. | MSW for waste generation: businesses and households, and for waste treatments waste treatments plants | This information allows calculation of the ratio between the MSW generated (by households and enterprises) and the different treatments, for example the share of recycled material on the total generated. Also the share of landfilled waste could be important: if this share is high, maybe not a large amount of waste collected is differentiated. It is important to stress the presence of detailed data by NUTs 2 level (region) | http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=en_v_rwas_gen&lang=en |

| Description | Unit of measure | Area geo | Statistical unit | Connections with public innovation | Link |
|--|--|---|--|---|---|
| Collection, reuse and recycling, and recovery of waste electrical and electronic equipment (WEEE) | Waste generated and treated per year kg per capita | EU Member State and EEA/EFTA countries. | Electrical and electronic equipment by all sectors and households. | Data seems to be useful in order to consider the category of e-waste (related to ICT) data concerns total waste collected by households and the amounts of e-waste gives a possible measure of e-waste from MSW | http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=en_v_waselee&lang=en |
| This dataset shows data provided by Member States' authorities through the questionnaire on the state of the environment (OECD/Eurostat). This dataset presents trends in amounts of municipal (including household waste), and the treatment and disposal method used. Furthermore, the amount of waste generated in each country is related to the rate of urbanisation, the types and pattern of consumption, household revenue and lifestyles. | Thousand tons | OECD country level | MSW generated and treated | Data could be useful in order to analyse the share of MSW generated with respect to different kind of treatment | https://stats.oecd.org/Index.aspx?DataSetCode=MUNW# |
| The rate of separate collection of municipal waste in Italian regions from 2010 to 2014 | Percentage | Italy | Italy, separate collection of municipal waste by region | Data give information about the different level of waste collected in different Italian regions (maybe information about how the service works: efficient or not across regions). | ISPRA Report |
| The amount of packaging collected by waste fractions (glass, wood, metals, etc.) | 1 000* tons | Italy | Separate collection of municipal packaging waste | Waste differentiated collection considering only the packaging | https://www.eionet.europa.eu |
| Costs distribution about waste management | eur/inhabitant | Italy | Costs analysis in function of MSW management | The cost distribution could provide a sort of proxy of level of service innovation | http://www.isprambiente.gov.it/it/archivio/eventi/2017/ottobre/rapporto-rifiuti-urbani-edizione-2017 |

| Description | Unit of measure | Area geo | Statistical unit | Connections with public innovation | Link |
|--|--|-------------|-------------------------------|---|---|
| The amount of waste generated and differentiated by regions, provinces and municipalities | Tons and Kg/inhabitants | Italy | MSW collected and generated | The ratio between the amount of waste generated and differentiated is a possible measure of the service efficiency | http://www.catasto-rifiuti.isprambiente.it/ |
| Household waste refers to waste from households collected by or on behalf of municipalities. Waste from small shops, etc. is sometimes collected together with household waste, hence a (small) part of household waste does not originate from households (Statistics Netherlands, 2015). | 1 000*tons | Netherlands | MSW by households | The share of different fraction of waste collected. This data combined with MSW generation give information about the efficiency of the systems and if some fractions of waste collected works better. | https://www.eea.europa.eu/publications/municipal-waste-management-across-european-countries |
| Total collected amount of fine residual waste (kg per inhabitant), per municipality | kg/inhabitant | Netherlands | MSW collected and generated | Small amounts of fine residual waste are a proxy for the overall performance of municipalities with regard to waste | https://afvalmonitor.databank.nl//Jive |
| Total separated amount of plastics, cans, drinks carton at sorting installation (kg per inhabitant), per municipality | kg/inhabitant | Netherlands | MSW collected and generated | Relatively large amounts of plastics, cans, drinks cartons are a proxy for the overall performance of municipalities with regard to waste. Should be benchmarked against total amount of fine domestic waste generated. | https://afvalmonitor.databank.nl//Jive |
| Share of domestic waste that is separated (e.g. into biological waste, paper, glass) and that can be reused | % | Netherlands | MSW collected and generated | Proxy for the overall performance of municipalities with regard to waste | https://afvalmonitor.databank.nl//Jive |
| Type of tariff system that is being used in a municipality (e.g. fixed, based on volume, etc.) | Five classes: fixed tariff, number of persons in household, weight, volume; volume and frequency | Netherlands | Municipal waste tariff system | Directly related to the scores in the web crawling pilot (e.g. a tariff based on weight is the most advanced level). | https://afvalmonitor.databank.nl//Jive |

| Description | Unit of measure | Area geo | Statistical unit | Connections with public innovation | Link |
|---|--|-------------|-------------------------------|--|---|
| Waste collection system that is being used by a municipality (e.g. minicontainers or bags). | Three classes: bags, minicontainers, duocontainers | Netherlands | Municipal waste tariff system | Directly related to the scores in the web crawling pilot (e.g. containers can be equipped with a chip, bags cannot). | https://afvalmonitor.databank.nl//Jive |
| Subcategory of overall sustainability score with special reference to waste and raw materials. Also includes indicators on the themes soil, water, air, pollution, nature and landscape, energy. For a detailed overview see: Telos (2017). Nationale Monitor Duurzame Gemeenten 2017. Telos: Tilburg (Table B.2, pp. 145-148). | Index score {0-100} | Netherlands | Municipalities | Gives a first impression of the performance of individual municipalities with reference to sustainability | http://www.sustainablecitiesbenchmark.eu |

Source: StarPIN data.

Table 9: A simplified template structure to integrate data on public sector innovation in e.g. the waste management domain

| Data collected on innovation in the public sector/public services | | | | | | | | | | | |
|---|----------|----------|------------------|--------------------------------------|-----------------------------|---|--|-------------------------------------|-----|-----|----------------|
| Country | Region | Province | Statistical code | Geographical/ institutional level | New data | Administrative data | | Statistical data connected with PSI | | | Features |
| | | | | Municipality | Indicator of innovativeness | ... from MUD (Unified environment statement form) | | % of differentiated on total waste | ... | ... | No inhabitants |
| Country X | | | | | 2 | | | 66 | | | 60 000 000 |
| Country X | Region Y | | | | 2 | | | 75 | | | 7 000 000 |
| Country X | Region Y | AA | | Municipality a | 1 | | | | | | 800 000 |
| Country X | Region Y | BB | | Municipality b | 3 | | | | | | 50 000 |
| Country X | Region Y | CC | | Municipality c | 2 | | | | | | 60 000 |
| Country X | Region Z | | | | 1 | | | 40 | | | 4 000 000 |
| Country X | Region Z | DD | A999 | Municipality d | 2 | | | | | | 100 000 |
| Country W | ... | | | | 3 | | | | | | 10 000 000 |
| | ... | | | ... | | | | | | | |

Source: StarPIN data.

The tables illustrate how data can be ordered by the corresponding geographical and institutional level: data referred to municipalities in the corresponding municipality line, data referred to the region or country level to be compared with a corresponding indicator value which will be calculated for the region or country on the basis of the score detected at the municipality level.

Pilot service Health homecare

In this section, a methodological approach to design and testing a new innovativeness indicator for the second pilot service domain, i.e. Health homecare is described.

Some methodological issues

Innovative levels in homecare

Homecare can be identified at the intersection of COFOG-CPA classifications (COFOG 07.02 Outpatient services; CPA 86.21.10 General medical practice services). Indeed, the notion of homecare as a public service relevant for innovation is undeniable. However, homecare can be a generic public service available for everyone, but it is quite often conceived for clearly identified categories of citizens (e.g. elderly people, children, people with specific disabilities, etc.).

The lack of clear boundaries and the plurality and heterogeneity of potential users in this service domain raises several methodological difficulties: the unit of analysis is not easy to identify, the value to be created differs according to user categories, actors involved in innovation vary across application areas. However, assessing innovation in this domain is particularly relevant because:

- i. Homecare is a public service needing strong innovation rates, especially for lowering the costs of healthcare in Europe, or for making public resources more efficient. The increasing age of European citizens reinforces this need for innovativeness in this public service;
- ii. Homecare is a public service where the adoption of innovation actually takes place, also thanks to the growing diffusion of ICT in society as well as to the digitalisation of public administration;
- iii. Despite being more relevant for specific categories of citizens, innovations in homecare can easily benefit the whole society, in two ways: enabling citizens to gain access to services directly from their home reduces the costs of mobility and of services, regardless of how difficult it is for patients to reach a hospital; resources saved by developing more efficient homecare can be devoted to increasing the resources for traditional healthcare (e.g. for hospitals).

In identifying the most appropriate levels of innovativeness, one problem has been due to the variety of keywords to be assigned at each level. If the baseline is easily attributable to public bodies that do not provide any type of homecare assistance on the territory, **level 1** is identified with the 'traditional' homecare. Traditional homecare is defined as 'supportive care provided in the patients' home by healthcare professionals'. Thus, the analysis focuses on healthcare professionals — such as nurses or physiotherapists — providing care at patients' homes. This is the case for patients that cannot, temporarily or permanently, go to a hospital to receive that service and therefore ask for its provision at home. Public bodies usually identify a selection of services to be included in the homecare list. However, the availability of the homecare service does not depend on technological issues, rather on the scarce resources of public bodies in charge of healthcare services in that territory.

Moving to the **level 2** of innovativeness, an additional issue can be raised. The concept here is 'care at distance' and, more technically, of tele-medicine. Since the 1970s, tele-medicine has been defined

in very different ways. The current WHO definition is: ‘the delivery of healthcare services, where distance is a critical factor, by all healthcare professionals using information and communication technologies (ICT) for the exchange of valid information for diagnosis, treatment and prevention of diseases’. From this perspective, healthcare professionals can monitor — albeit partially — their patients at a distance. Over the years, several instruments and devices have been produced with this scope, such as for instance: examination cameras, interactive stethoscopes, digital ECGs and so on.

The identification of the highest level of innovativeness (**level 3**) is more challenging. A decision was taken to refer to this level as the one most characterised by web-based provision of home care services (eHealth). However, a univocal definition of eHealth is still missing. WHO defines eHealth as the use of information and communication technologies (ICT) for health. That is almost overlapping with tele-medicine, which is our level 2. In other cases, eHealth is even broader than tele-medicine, including for instance the delivery of medical reports or even reservation of medical examinations via telematics. This definition was therefore too broad for our homecare case.

The objective targeted with level 3 was the use of ICT tools for homecare purposes. Devices like smartphones or tablets, which are likely to substitute dedicated homecare/biomedical devices, have been taken into consideration, possibly jointly with the use of applications (apps). This is the concept of mobile Health (mHealth): ‘use of mobile wireless technologies for public health, reflecting the increasing importance of this resource for health services delivery and public health, given their ease of use, broad reach and wide acceptance’ (WHO, 139th Executive Board, 2016, Geneva, Switzerland). Of course, in this study, the concept of mHealth refers only to homecare services.

Summing up, ‘traditional’, ‘tele*/distance’ and ‘mobile’ concepts are all generic innovative levels that can all be applied to care of patients at their home (homecare), as in our pilot case.

Identification of keywords

In order to assign the innovative level to each homecare service, several keywords were selected. At first, it was decided to assign to level 1 all words concerning traditional homecare, including words like nurse, physiotherapist or similar. In level 2, all words concerning the ‘at distance’ concept were included, thus listing specific homecare services like: tele-cardiology, tele-radiology, tele-diabetology and so on. Finally, level 3 was identified by words concerning eHealth, using mostly words coming from eHealth projects (e.g. EU-funded research projects or pilot projects by health authorities). However, this strategy was not fully successful, so in a second round it was decided to split innovative levels and keywords as follows:

Level 1

- Basket 1 is the baseline and identifies all keywords referring to ‘care’, ‘health’, ‘assistance’ and so on;
- Basket 2 introduces words concerning the ‘at home’ concept: ‘homecare’, ‘domestic’, ‘domiciliary’ and so on.
- As a result, keywords lead to Level 1 whenever ‘healthcare’ and ‘at home’ concepts are found together.

Level 2

- Basket 3 introduces the ‘at distance’ concept with generic words like ‘distance’ itself or ‘tele-medicine’. All declinations of this word have been tried, such as for example: telemedicine, tele medicine, tele-medicine.
- Basket 4 replicates this idea but with a long list of ‘tele*’ services (such as mentioned before, tele-cardiology, tele-radiology, etc.). All declinations of these words have been tried, such as for example: telecardiology, tele cardiology, tele-cardiology.

- As a result, keywords lead to Level 2 whenever the ‘healthcare’ concepts are found together with the ‘at distance’ concept or with one of the specific tele* services, grouped in basket 5.

Level 3

- Finally, level 3 is identified whenever, together with words under basket 1 and one among baskets 4 or 5, also keywords of basket 6 like self-positioning tools or basket 7 like interactive, bidirectional tools including ‘smartphone’, ‘tablet’, ‘apps’ or similar are found.
- As an additional note, this pilot case has raised the issue of whether mHealth is changing the concept of this public service. That is, mHealth does not imply receiving care ‘at home’ but wherever; this might mean neglecting the concept of the ‘home’ care itself. In order to prevent any confusion about the actual object of the analysis, which would affect the whole pilot case, it was assumed to consider mHealth a simple shift in the technology used for homecare services, as stated, using apps and smartphones/tablets instead of dedicated bio-medical devices.

Identification of units of observation

Differently from other cases, the healthcare/homecare service implies another methodological issue. That is, the identification of the most relevant unit of observation, dealing with the provision of health services. In different countries, in fact, very different institutions might be in charge of homecare services. Similarly, the territorial dimension can be very different, with EU countries keeping the provision of healthcare services at the national level and others breaking them down across regions. As a result, homecare services might be provided by dedicated health agencies (such as ASL in Italy), by territorial administrations (regions, municipalities) or by hospitals directly. From the patient’s point of view, then, the perception of the providing subject can also be very different from the subject that is responsible for the service. Patients perhaps directly contact their municipality, also because homecare services are often included in the more generic social assistance. However, assigning the innovativeness level to municipalities would be misleading, as many other social assistance services might be found, with no univocal reference to homecare.

Overall, the territorial dimension has to be taken into great consideration. Even if ‘at distance’, homecare services would always be provided by (or on behalf of) public subjects relatively close to patients. Thus, for this pilot case, it was decided to go down to the other two most disaggregated levels: health agencies and hospitals.

The validation process

Results from web-scraping in the homecare pilot have been checked qualitatively and also compared with available administrative data. The main methodological issues are discussed hereafter.

1. Missing units of observations

In the Italian health agencies’ case, web-scraping results concern 89 agencies, while the web-scraping was launched on 107 websites (83 %). These websites had been identified on an Italian administrative database, thus it was expected to get a 100 % coverage. However, this database might be outdated. In fact, the numbers of Italian territorial health agencies have been constantly decreasing over the last 25 years (see Table 9). In another publicly available source it is reported that in year 2016 they were 121 (a higher number than that used for the web-scraping), while in 2017 just 101 (thus a lower number). It is probable that some health agencies have been merged together between 2016 and 2017, with the result of changing the website address. If the administrative database has not been modified accordingly, the web-scraping could gather data on a maximum number of 101 agencies. This brings our coverage up to 88 % (89 websites scraped out of 101).

Table 10: Number of health agencies in Italy

| Year | 1992 | 1995 | 2004 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 |
|---------------------------------|------|------|------|------|------|------|------|------|------|------|
| Number of health agencies (ASL) | 659 | 228 | 183 | 145 | 145 | 143 | 140 | 139 | 121 | 101 |

Source: StarPIN data.

Concerning the missing 12 %, qualitative checks have put forward the following motivations:

- i. Automatic redirection to a different spelling of websites. Examples are:
 - <http://portale.asl.at.it> redirects to
 - <http://portale.asl.at.it/Apps/portaleasl.nsf/index.htm>
 - <http://www.aslnapoli3sud.it> redirects to
 - <http://www.aslnapoli3sud.it/sitoweb/jportal/JPMMain.do>
- ii. Radical change in the official name of the health agency and, consequently, of the website address. This is the case with <http://www.usl12.toscana.it> which is now <http://www.uslnordovest.toscana.it>.

2. Low number of observations (web pages)

For some units (health agencies), there is a very low number of observations (web pages). In some cases, the web scraping stops after scraping the homepage. In other cases, it stops after scraping a very limited number of webpages (between 2 and 7). For analytical purposes, it is suggested to delete these observations as they might imply a bias.

Possible motivations of this problem are:

- i. Automatic redirection not properly working in the web-scraping, due to technical reasons (e.g. old name of <http://www.ass3.sanita.fvg.it>); although there is an automatic redirection to the new website address, the web-scraping stops.
- ii. Other technical motivations as highlighted previously in this report.

From an analytical point of view, this problem leads to the definition of an innovativeness level equal to 0, that simply reflects the missed web-scraping on the entire websites.

3. False positive cases

A qualitative check of the web-scraping results has also allowed the identification of ‘false positive’ cases.

Some of these cases — e.g. the one referred to the Italian health agency <http://www.asp.crotone.it> — are due to overabundance of web pages. This is the opposite of the previous methodological problem: some health agencies have websites with a very high number of pages (the one mentioned here has 476 active web pages on which the web-scraping has been performed; other health agencies have websites with more than 1 000 web pages). Such overabundance might imply a large variety of words — including keywords for assessing the innovativeness level — thus distorting the probability to match words uniquely. The result is a high correlation of number of web pages and level 3 of innovativeness. The methodological suggestion, in parallel to the opposite case highlighted before, is to carefully and qualitatively check the results of those units of observations with a very high number of web pages.

Also in the case of healthcare, web-scraping can yield valuable data to identify innovation levels, to be validated by means of other complementary, control data. Table 11 provides a description of some of the complementary statistical and administrative sources to be used in such a validation process.

Table 11: Description of data available on and related to home care from different sources, for the EU and Italy

| Description | Unit of measure | Area Geo | Statistical unit | Connections with public innovation | Link |
|---|--------------------------|------------------|--|--|--|
| <i>Data on the access to professional homecare, across EU Member States.</i> | Percentage of population | EU Member States | The indicator covers persons paying professional homecare services by household type, income group and level of difficulty to pay professional homecare services | Data give information about the relevance of homecare services in EU countries. Lower levels of this indicator are expected to be more correlated with lower levels of the innovation indicator. | http://appsso.eurostat.ec.europa.eu variable 'ilc_at16' |
| Data on the access to professional homecare, across EU Member States; including motivations for not accessing homecare (Refused by person needing such services; vs. Financial reasons) | Percentage of population | EU Member States | The indicator covers persons using or not using professional homecare services by household type, income group, degree of urbanisation and reason for not using professional homecare services | Also these data report on the relevance of homecare services, linking it to the motivations. It can be expected that lower levels of innovativeness correlate with a higher level of this indicator under the motivation 'financial reason'. Lower income countries are less likely to invest in higher innovativeness levels of this service. | http://appsso.eurostat.ec.europa.eu variable 'ilc_at15' |

| Description | Unit of measure | Area Geo | Statistical unit | Connections with public innovation | Link |
|--|---|---|---|---|---|
| Catholic University 'Rapporto Osservasalute'. It is an observatory on healthcare across Italian regions. Includes the indicator: integrated home healthcare. | Percentage of population getting access to homecare; percentage of ASL providing homecare services. | Italian NUTS2 regions | Both population and health agencies | Aggregating data at the regional level, it might be possible to double check data on innovativeness levels, e.g. if higher shares of health agencies providing homecare in a territory (region) correlate with higher average values of innovativeness in that same territory | http://www.osservatoriosullasalute.it |
| List of public entities reporting 'homecare data' (time series) | Percentage of public entities | Italy (potentially all EU Member States) | Public entities (e.g. health agencies, hospitals) | It can constitute a rough measure of consistency: the absence of a specific public entity is expected for those entities with innovativeness level equal to 0. Similarly, entities with a higher level of innovativeness are expected to be found more than one time (e.g. once for a 'traditional homecare service' eventually provided in the past, and once for a tele-medicine service) | http://basidati.agid.gov.it/catalogo/index.html |
| List of public (and private) structures providing homecare services. Within the Lombardy region open data | Percentage of public entities | Lombardy region (Italy) and potentially other regions | Public entities (e.g. health agencies, hospitals) | It is a database that allows to qualitatively check the levels of innovativeness of public entities within that specific region | https://www.dati.lombardia.it |
| Italian Health Ministry SIAD (Informative system on homecare) | - | Italy | Public entities (e.g. health agencies, hospitals) | Ministry database covering all homecare services providers. It might constitute the basis for several indicators on homecare. Unfortunately, its implementation is still pending | http://www.salute.gov.it/portale/temi/p2_5.jsp?lingua=italiano&area=sistemaInformativo&menu=domiciliare |

Source: StarPIN data.

7

Main results of the methodology application

This chapter will present the main results of the pilot application of the StarPIN framework and methodology to five EU countries, with illustrative analyses of newly collected data. The bulk of data collection and elaborations refer to the Waste management service domain, built and refined in two pilot countries in particular, Italy and the Netherlands, and extended to Portugal, Slovakia and Sweden. Although with a lesser depth, the methodology has been applied also to another service domain, Health homecare, and results are shown with reference to Italy, the Netherlands and Portugal. The validation process and the cross-sectional analysis of the newly created data, together with administrative data and statistical data, which are connected to the public service innovation phenomenon and to the two pilot services in particular, are discussed from the methodological point of view for paving the way for future steps of the methodology application and for research work.

Pilot service waste collection in the five pilot countries

For waste collection, results for the innovativeness level of waste collection public service are provided for all cities with more than 50 000 inhabitants for the five pilot countries as test of the application of the methodology. The Dutch municipalities have been fully included in the first wave of web scraping. The much higher number of Italian municipalities was first tested in a small subset, and then — given also the poor results extracted from the websites of the very small municipalities — a priority was given to the selection of all the cities and towns which are capital of a Province (NUTS2) and of other municipalities with more than 50 000 inhabitants. For the other three countries, the web-scraping proceeded in two steps: first, a very small subset illustrated in the following table; second, again starting from the larger cities/towns in terms of inhabitants, the coverage was brought up to a number of municipalities including from one fourth to one third of the national population.

The most aggregate result obtained for the first pilot service examined is a country-level innovativeness indicator for waste collection public service for the five pilot countries: 2 275 for Italy, 2 795 for the Netherlands, 2 134 for Portugal, 2 126 for Slovakia and 2 997 for Sweden.

A preliminary test in Portugal, Slovakia and Sweden

After completing the data collection for Italy and the Netherlands, the preliminary test on the three additional countries has been conducted on 20 randomly extracted municipalities for each country. Website-centric data collection yielded results for 18 of these municipalities in Sweden, 16 in Portugal and 13 in Slovakia.

These very preliminary results are shown in the following table to share the first image of the permeability and of the polarisations recorded by a limited number of observations.

Table 12: Test for waste collection service innovativeness level for the first 20 municipalities in the three additional pilot countries: Portugal, Slovakia, Sweden

| Innovativeness level | Portugal | | Slovakia | | Sweden | |
|----------------------|----------------------|------------------|----------------------|------------------|----------------------|------------------|
| | No of municipalities | % municipalities | No of municipalities | % municipalities | No of municipalities | % municipalities |
| 0 | 2 | 12.50 % | | | 1 | 5.60 % |
| 1 | 3 | 18.80 % | 1 | 7.70 % | | |
| 2 | 5 | 31.30 % | | | 2 | 11.10 % |
| 3 | 5 | 31.30 % | 12 | 92.30 % | 13 | 72.20 % |
| 4 | 1 | 6.30 % | | | 2 | 11.10 % |
| Total | 16 | 100.00 % | 13 | 100.00 % | 18 | 100.00 % |

Source: StarPIN based on web scraping.

The data shown in Table 12 highlight a substantial heterogeneity in innovation levels, which appears to be more evenly distributed across innovation levels in the case of Portugal, concentrated in the medium-high level in the case of Slovakia, and with the highest frequency of municipalities with the top innovativeness in Sweden. However, given the very limited number of websites visited, the test was used mainly to check the possibility to access the website, or rather the potential presence of obstacles of various nature.

Data collection extended to all five pilot countries

The full application of the web-scraping methodology to the pilot service of municipal solid waste collection allowed assigning a specific innovativeness level to each municipality. A synthetic picture of innovativeness levels for waste collection in the five countries is summarised in the following tables.

As mentioned already, the methodology is at its first pilot application, results must therefore be considered quite preliminary, still data are illustrative of the type of information that could be collected, and some tentative considerations can be made.

A caveat applies here: Any cross-country analysis of innovativeness levels would require that data are normalised to account for the substantial diversity of the municipalities under observation, and of the socioeconomic characteristics of countries and regions in which they are located. Some complementary data on waste collection at different levels of geographical aggregation that could be used to contextualise the analysis of innovation data are available in Tables 14, 15 and 16.

Table 13: A synthetic view of the waste collection web-scraping data in five countries, June 2018

| | Italy | Netherlands | Portugal | Slovakia | Sweden |
|---|--------------|--------------------|-----------------|-----------------|---------------|
| Nr. municipalities visited | 328 | 324 | 99 | 28 | 40 |
| Country size in terms of population (thousands) | 60 484 | 17 020 | 10 320 | 5 429 | 9 903 |
| % population included in the visited municipalities | 36.37 | 86.60 | 68.54 | 26.24 | 35.47 |
| % of municipality waste collection attributed to each service innovativeness level | | | | | |
| Innovativeness level | IT | NL | PT | SK | SE |
| 0 | 9.45 % | 0.61 % | 20.20 % | 0 | 5.00 % |
| 1 | 29.27 % | 11.28 % | 29.29 % | 28.57 % | 2.50 % |
| 2 | 31.40 % | 10.06 % | 15.15 % | 3.57 % | 2.50 % |
| 3 | 27.44 % | 65.24 % | 33.33 % | 67.86 % | 77.50 % |
| 4 | 2.44 % | 12.80 % | 2.02 % | 0 | 12.50 % |

Source: StarPIN based on web scraping.

Still, a few comments can be made on these data. First, the coverage of population is quite different in the five countries. The average size of municipalities is generally quite different across the five countries. Approximately the same number of municipalities corresponds to 86 % of the population in the Netherlands and only one third of inhabitants in Italy, due to a substantial diversity in both the institutional structure and the size of the two countries. The coverage of population is much more similar when comparing Italy to Slovakia or to Sweden.

Second, the percentage of websites exhibiting a value of the indicator equal to zero is much higher in the case of Italy and Portugal. It is quite likely that these high percentages reveal that websites could not be accessed for some technical reasons. This is a problematic issue that needs to be explored in depth to improve the quality of the data collected (see Chapter 4 for a discussion on these data accessibility problems, and how they can be tackled).

Third, the distribution of innovativeness levels differs remarkably across countries. Level 3 is more than twice as high in the case of the Netherlands, Slovakia and Sweden as compared to the case of Italy and Portugal. Sweden and the Netherlands exhibit the highest share of municipalities with the top level of innovativeness. The low score of Italian municipalities is even more striking, considering that our sample excludes small-sized towns (below 50 000 inhabitants) which represent the bulk of urban centres in Italy.

Tables 14 through 16 provide examples of data that are complementary to the web-scraping and can be used to validate the web-scraping-based indicators.

Table 14: Data on waste collection in the Netherlands (see the complete table in Annex 4)

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|-------------------|--|-------------------------|-------------------|---------------|---------------|----------------------|
| Pekela | Number of residents | | | | | |
| Druuten | Expensive bags and number of residents | Minicontainers | 49 | 0.86 | Every 2 weeks | 72 |
| Steenwijkerland | Volume AND frequency | Minicontainers | 68 | 0.82 | Monthly | 70 |
| Bronckhorst | Volume AND frequency | Minicontainers | 81 | 0.78 | Monthly | 70 |
| Beuningen | Expensive bags AND | Minicontainers AND bags | 90 | 0.77 | Every 2 weeks | 69 |
| Heumen | Expensive bags and number of residents | Bags | 55 | 0.85 | Every 2 weeks | 69 |
| Eersel | Volume AND frequency | Minicontainers | 52 | 0.82 | Monthly | 69 |
| Lochem | Volume AND frequency | Minicontainers | 106 | 0.7 | Every 2 weeks | 69 |
| Berg en Dal | Different systems in parallel | Minicontainers AND bags | 0 | | Every 2 weeks | 68 |
| Heeze-Leende | Volume AND frequency | | 57 | 0.79 | | 68 |

Source: StarPIN based on web scraping, data processing of Datalogic for StarPIN.

Table 15: Data on waste collection in Portugal

| Name of municipality | Nr. of webpages | Waste collection service innovation level | District/ Autonomous region | Pop.total [c] | PD (per km ²) | Ruling party |
|----------------------|-----------------|---|-----------------------------|---------------|---------------------------|--------------|
| Chaves | 82 | 3 | Vila Real | 44 186 | 75 | PS |
| Águeda | 58 | 3 | Aveiro | 49 691 | 148 | IND |
| Albufeira | 121 | 3 | Faro | 35 281 | 251 | PSD |
| Alcobaça | 3 | 2 | Leiria | 55 269 | 135 | PSD |
| Almeirim | 21 | 3 | Santarém | 22 617 | 102 | PS |
| Amadora | 17 | 3 | Lisboa | 176 239 | 7 405 | PS |
| Amarante | 8 | 2 | Porto | 61 029 | 202 | PSD-CDS |
| Anadia | 21 | 3 | Aveiro | 31 671 | 146 | IND |
| Avis | 18 | 1 | Portalegre | 5 054 | 8 | PCP-PEV |
| Azambuja | 10 | 0 | Lisboa | 21 508 | 82 | PS |
| ... (cont.) | | | | | | |

Source: Pordata.pt.

Table 16: Data on the share of differentiated waste collected at the Municipality level — Italy

| Region | Capital of the province | Other municipalities | Inhabitants | Service innovativeness level | Scaling up country and region level | # webpages | Region differentiated % of collected waste 2016 | Province differentiated % of collected waste 2016 | Differentiated % of collected waste 2012 in the capital municipality of the province |
|----------------------|-------------------------|-----------------------|---------------|------------------------------|-------------------------------------|------------|---|---|--|
| Italy | - | - | - | - | 2 | - | - | - | 34.9 |
| Piemonte | - | - | - | - | 2 | - | 56.63 | - | - |
| | Torino | | 886 837 | 2 | | 9 | | 52.73 | 43.3 |
| | | Moncalieri | 57 530 | 0 | | 2 | | | |
| | Vercelli | | 46 552 | 2 | | 23 | | 63.91 | 67.7 |
| | | Civiasco | | 2 | | 25 | | | |
| | Novara | | 104 284 | 1 | | 10 | | 67.92 | 72.6 |
| | | Caltignaga | | 1 | | 24 | | | |
| | | Castellazzo novarese | | 1 | | 10 | | | |
| | Cuneo | | 56 124 | 2 | | 38 | | 60.69 | 43.9 |
| | | Alba | 31 453 | 3 | | 50 | | | |
| | | Monticello d'Alba | | 1 | | 30 | | | |
| | Asti | | 76 164 | 2 | | 35 | | 64.89 | 61.4 |
| | Alessandria | | 93 839 | 3 | | 6 | | 49.52 | 49.2 |
| | | Montemarzino | | 1 | | 27 | | | |
| | | Rivalta | | 2 | | 96 | | | |
| | Biella | | 44 616 | 2 | | 90 | | 61.36 | 56.0 |
| | | Cossato | | | | 1 | | | |
| | | Curino | | | | 14 | | | |
| | | Gaglianico | | | | 1 | | | |
| | Verbania | | | | | | | 66.85 | 73.0 |
| Valle d'Aosta | | | 34 361 | | 3 | | 55.60 | | |
| | Aosta | | | | | 41 | | 52.73 | 53.1 |
| | | Fenis | | 1 | | 12 | | | |
| | | Saint Rhemy en bosses | | 1 | | 61 | | | |
| Liguria | | | | | 3 | | 68.11 | | |

Source: StarPIN based on web scraping, ISPRA, ISTAT.

For each geographical/institutional level to which an innovativeness indicator can be attributed, administrative and statistical data available at the corresponding geographical level can be selected for future analyses. These analyses and the comparability between regions and countries strongly depend on the building of an indicator of innovativeness based on the observations made at a lower and smaller geographical level.

An additional step of the methodology consists of applying the method discussed in the following paragraph and described in Annex 7 for obtaining an innovativeness indicator for the country, and for the region when the observation is at the municipality or sub-regional level.

Scaling up the indicator generated

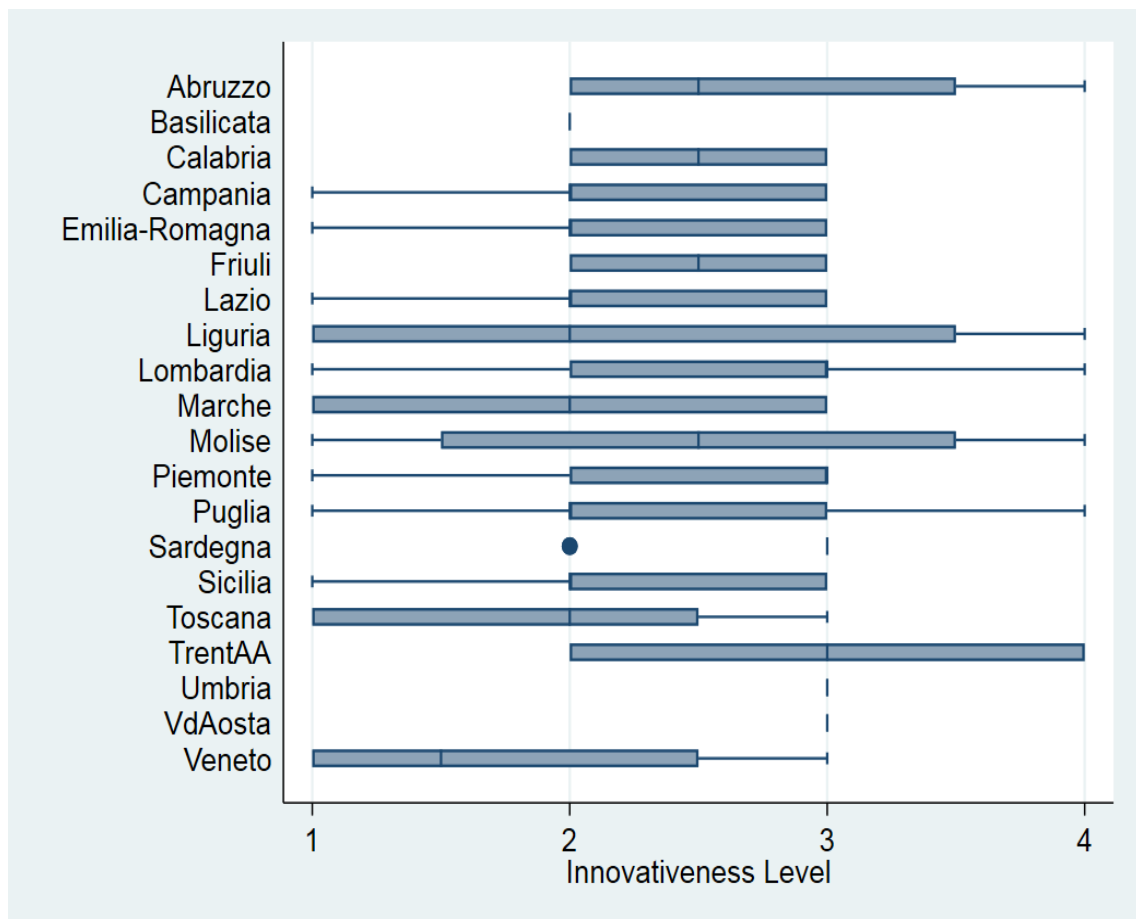
For comparative and analytical purposes, information on the innovativeness of individual subjects, now expressed at the municipality level in the case of waste management, may be usefully aggregated at a higher geographical (or sectoral) level.

An application of the weighted average calculation method described in Annex 5 for all five pilot Member States — allows for a first comparison between countries and highlights how subnational regions are distributed around the national means for each individual country.

Waste collection service innovativeness indicator, calculated as weighted average for the country, is 2 275 for Italy, 2 795 for the Netherlands, 2 134 for Portugal, 2 126 for Slovakia and 2 997 for Sweden.

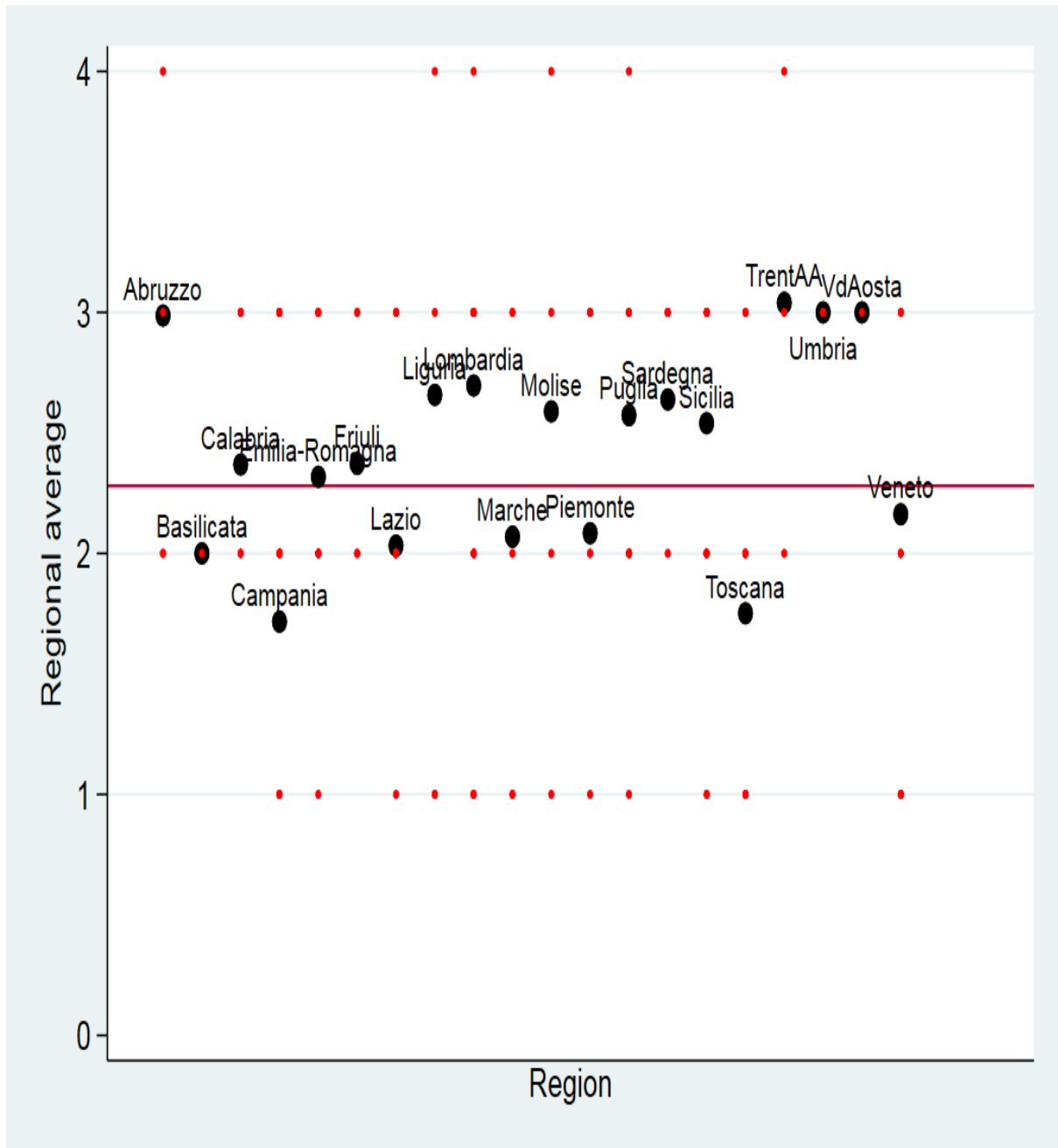
Both the box-and-whiskers plot and the distribution around the average provide an image of the range and variety of cases between regions within each country and between countries that was impossible to get by considering only the percentage of presence of municipalities in each level indicator class. In aggregate terms, the proximity to a higher level of innovativeness for waste collection service is confirmed for Sweden and the Netherlands, followed by Italy, Portugal and Slovakia at a similar level.

Figure 11: Weighted average score for Italian regions (StarPIN data)



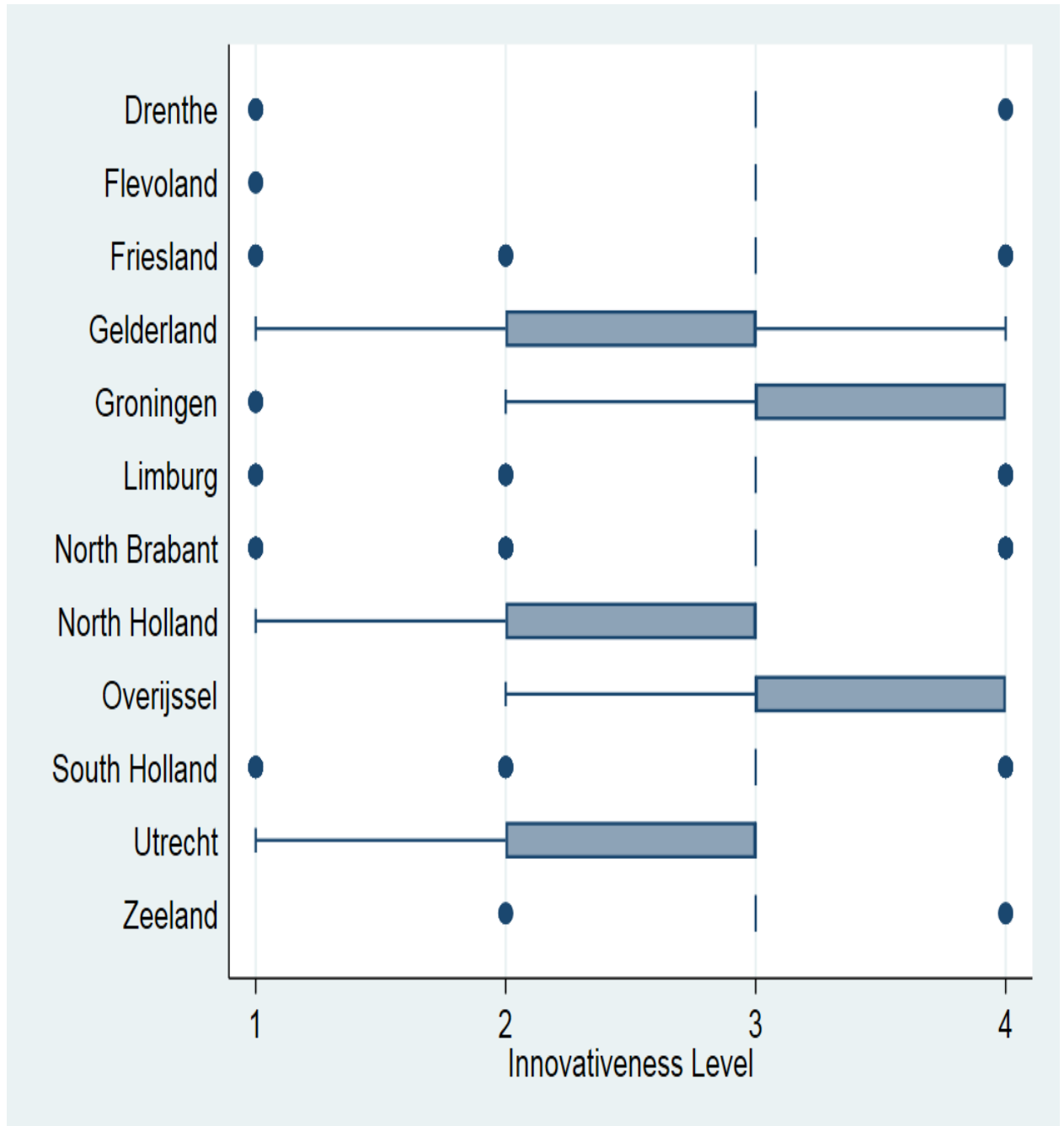
Source: StarPIN

Figure 12: Distribution of the 20 Italian regions around the average (StarPIN data)



Source: StarPIN

Figure 13: Weighted average score for Dutch regions (StarPIN data)



Source: StarPIN

Figure 14: Distribution of the 12 Dutch regions around the average (StarPIN data)

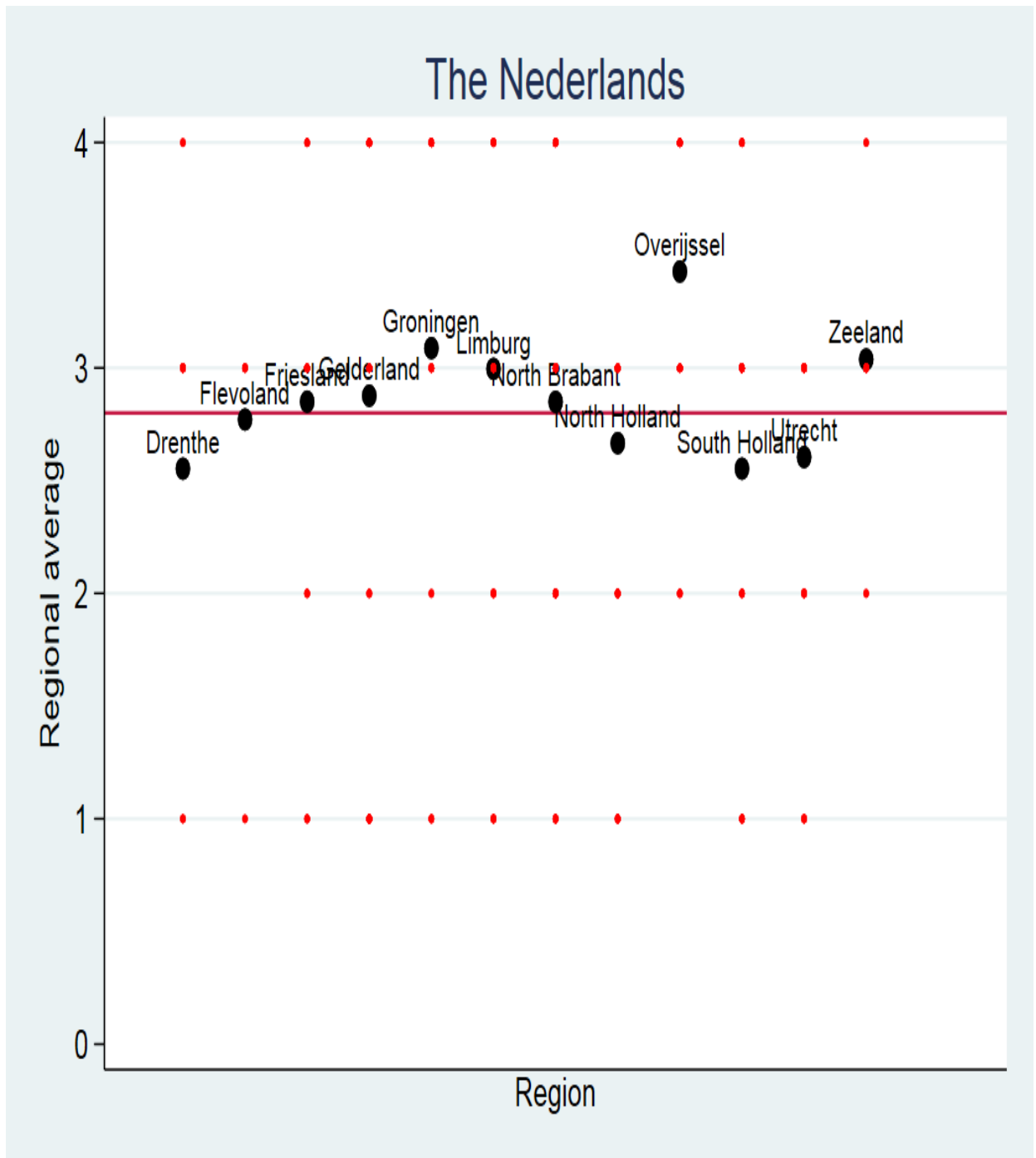
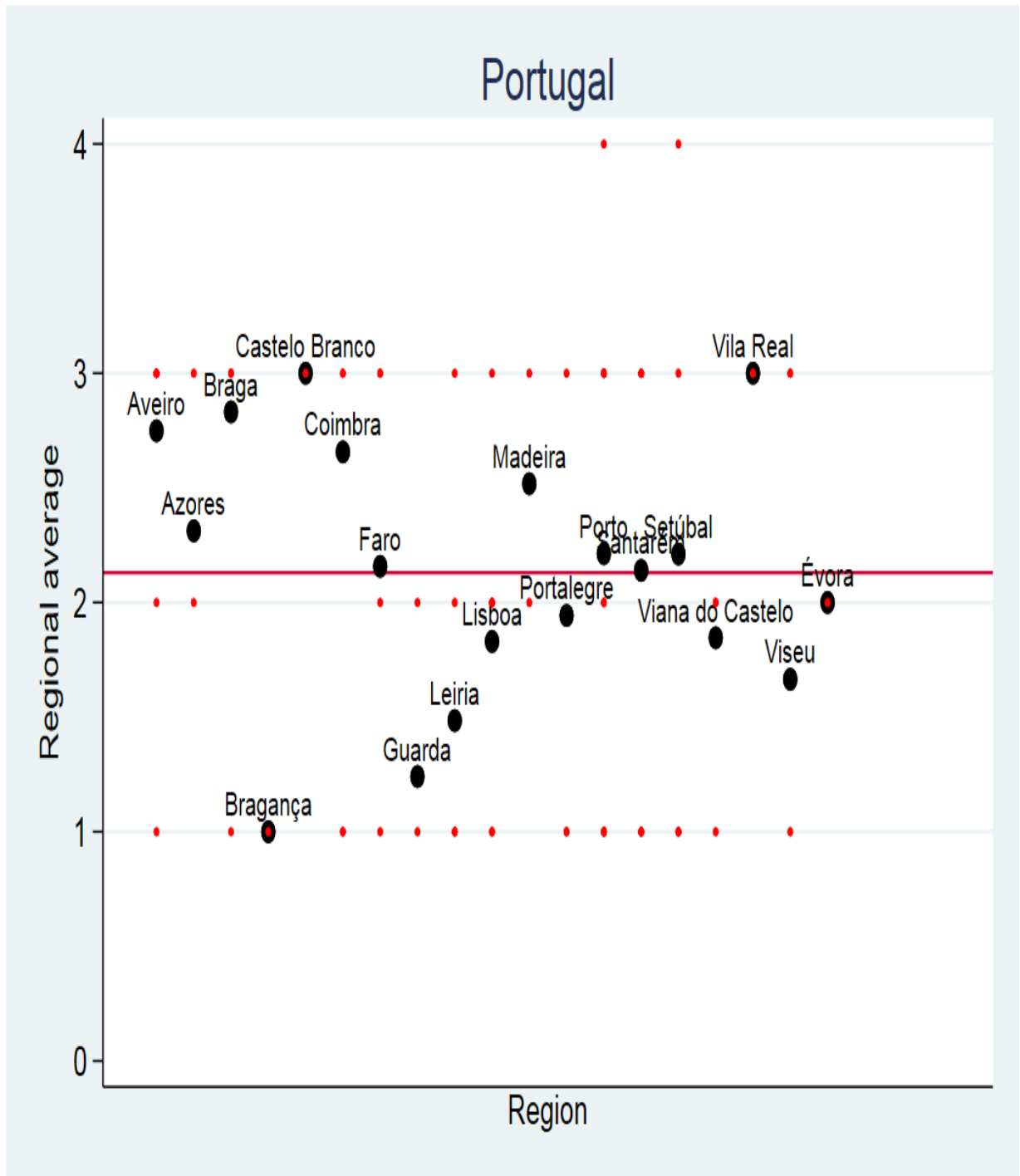
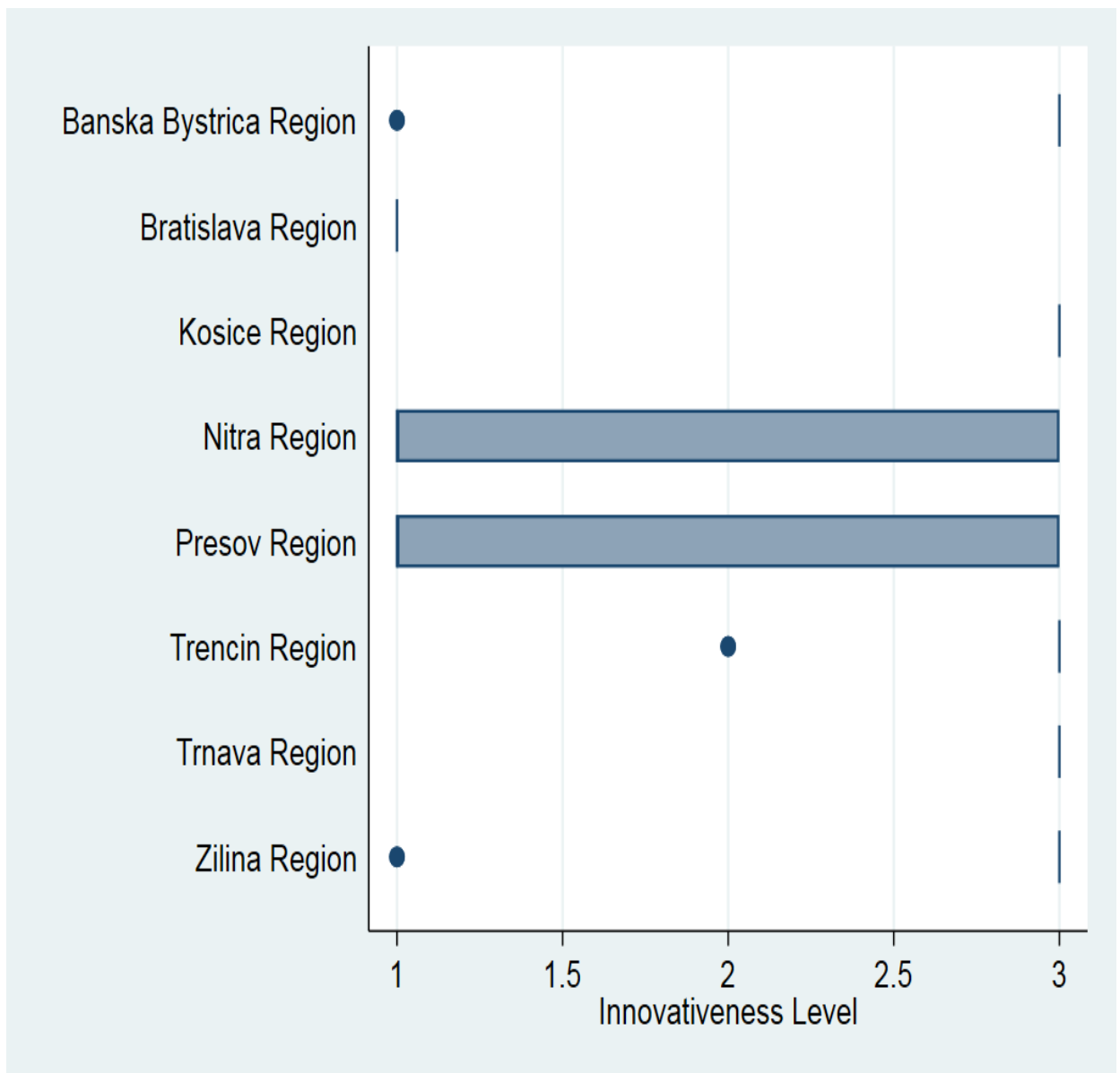


Figure 16: Distribution of the 19 Portuguese regions around the average (StarPIN data)



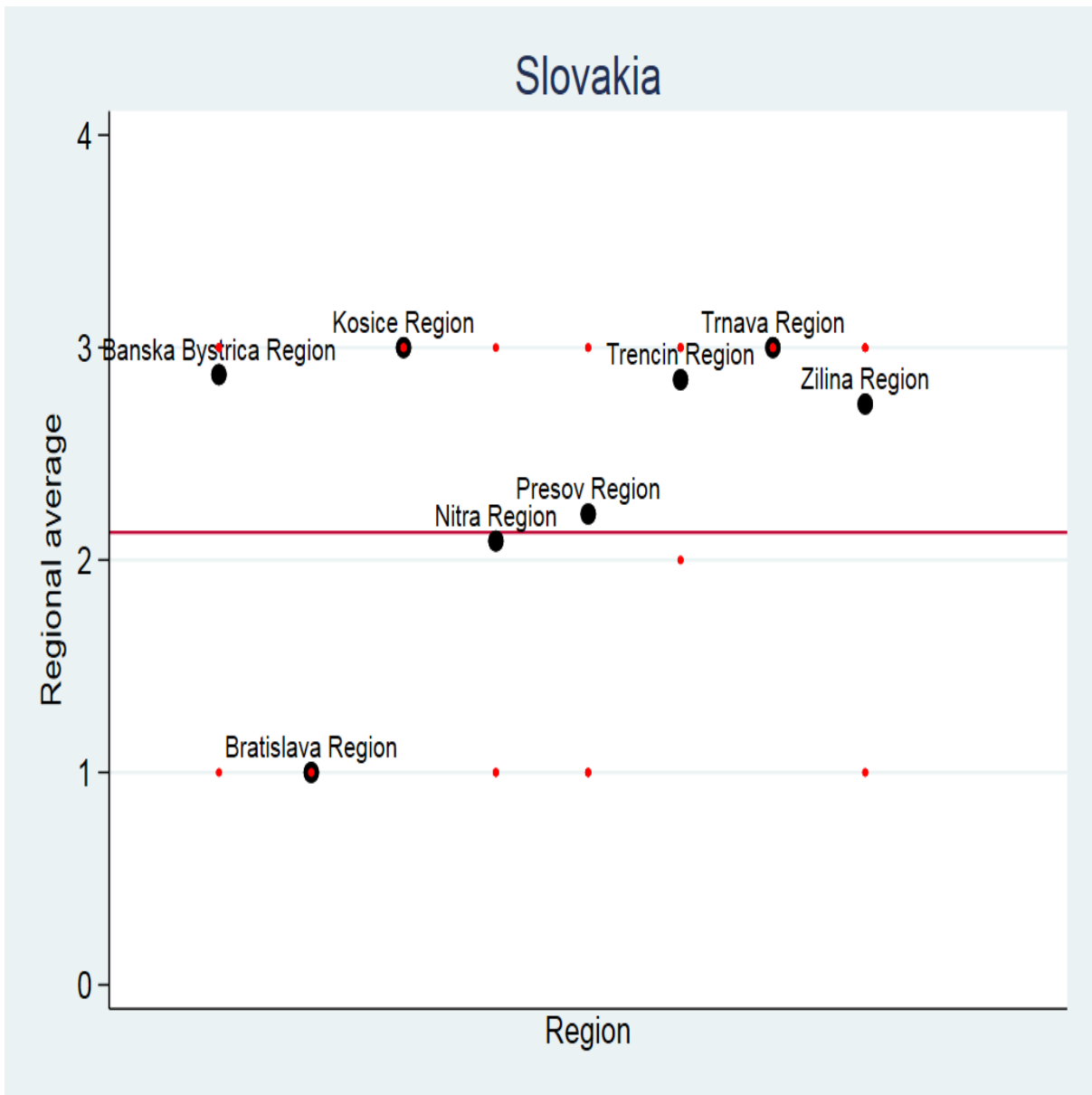
Source: StarPIN

Figure 17: Weighted average score for Slovakian regions (StarPIN data)



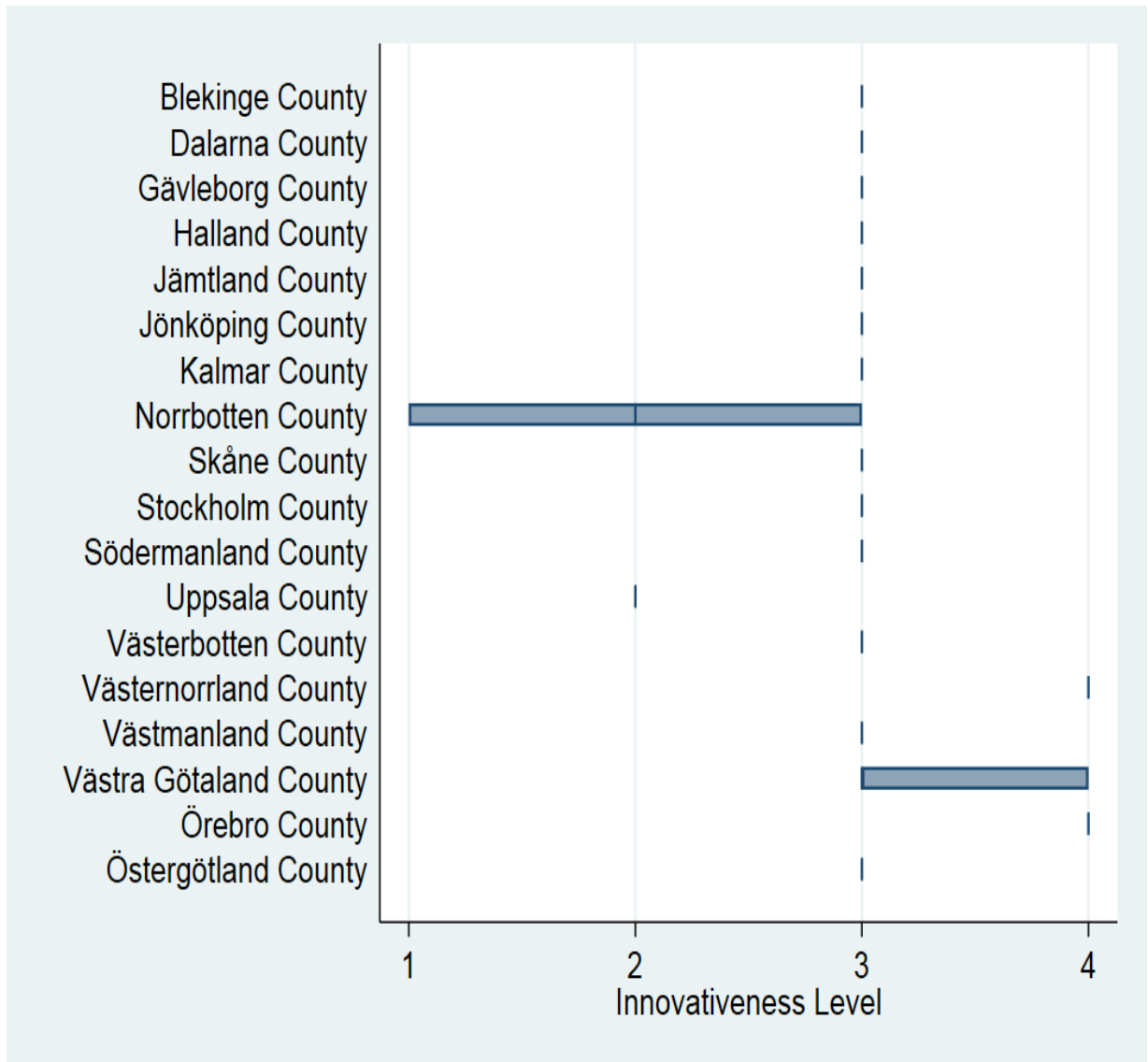
Source: StarPIN

Figure 18: Distribution of the eight Slovakian regions around the average (StarPIN data)



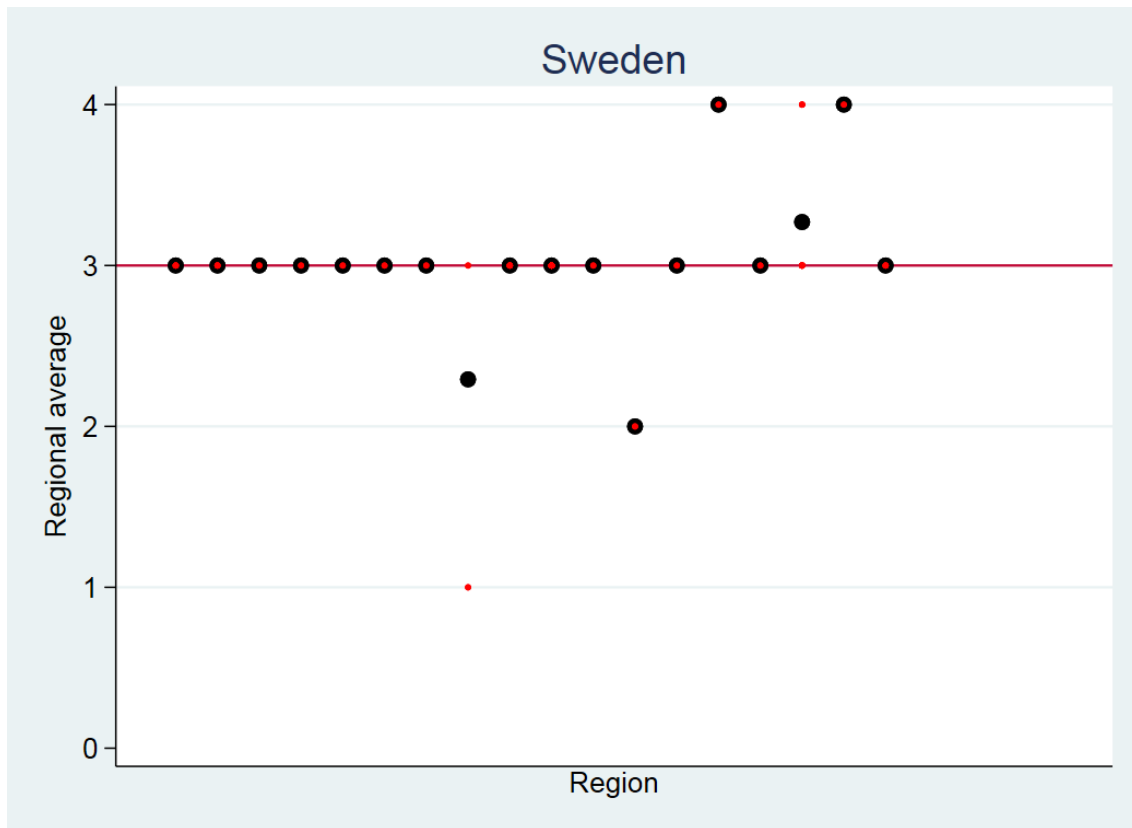
Source: StarPIN

Figure 19: Weighted average score for Swedish regions (StarPIN data)



Source: StarPIN

Figure 20: Distribution of the 18 Swedish regions around the average (StarPIN data)



Source: StarPIN

The indicator of innovativeness for the public service ‘waste collection’ described above classifies municipalities according to an ordinal variable. However, the innovation gap across different categories of the ordinal variable cannot be clearly measured. Differences in innovativeness reflect synthetic assessments of the public value created by different waste collection options: in some circumstances, the ranking of options could change depending on context-specific factors.

These considerations suggest that computing the simple or average value of the innovativeness recorded for different subjects within or across a given subnational context (e.g. cities, provinces or metropolitan areas) to scale up the indicator at the regional or national level gives some useful and synthetic information but may not be the most preferred solution. A first simple alternative to the average could be the simple or weighted median of the indicator. The simple median of the indicator indicates that half of the municipalities within a region (or a country) are below the value of the median, and half of the municipalities are above the value of the median. Alternatively, the median could be computed by attributing to municipalities weights that are proportional to the size of the municipality in terms of, for example, population or amount of waste generated. The weighted median indicates that half of the population (or waste) resides in municipalities with innovativeness below the median and half of the population (or waste) resides in municipalities with innovativeness above the median.

A more comprehensive approach could be that of considering the whole distribution of municipalities (and corresponding population or waste) across different levels of innovativeness. However, by considering six levels of innovativeness, this approach would result into six variables that need to be evaluated jointly. A more synthetic approach in this case would be to calculate the simple or weighted cumulative distribution of municipalities with an innovativeness index ‘greater than, or equal to’ a certain value.

These approaches to scaling up can be used only if information about the innovativeness level is

available for the population of municipalities within a region or country (e.g. the Dutch case). In some cases, however, information about innovativeness is only available for a selected sample of municipalities. In our pilot study, for example, the web-scraping tool is set to get information on the population of municipalities with 20 000 or 50 000 or more inhabitants, depending on the size of the country and its institutional structure. Depending on the administrative structure of the country of analysis, this threshold may allow to cover a very large or very small share of municipalities and corresponding population. This situation may impair the possibility to provide a representative picture of the innovativeness of a region or a country. Two alternative options are viable in this case. First, it is possible to change the population of reference for the measurement to match the sample that was selected. This approach is followed by many statistical offices who report statistics on large cities only for selected indicators (e.g. indicators that refer to urban phenomenon). Alternatively, it is possible to assume that the results obtained for the sub-sample of municipalities can be extended to the rest of municipalities that did not take part in the web-scraping. This latter approach, however, is prone to possible substantial biases, especially if the aim of the data collection is to compare different areas, while it may be a fair compromise if the objective is to perform some analytical work based on imperfect proxies.

Cross-sectional analysis of data

Country by country, service by service, and with different time spans, a variety of different correlation analyses will be possible. At the present stage of data collection, only cross-sectional analyses can be envisaged, connecting our innovation indicator to several characteristics of services, countries, regions, and individual institutions.

A non-exhaustive list of data categories that could be exploited to evaluate their correlation with public service innovation includes the following:

- i. Selected data at national and regional level Eurostat and OECD;
- ii. Data for one specific country at municipal level;
- iii. Administrative data — e.g. MUD ('Modello Unico di Dichiarazione Ambientale') for environmental services in Italy;
- iv. E-service data;
- v. ICT adoption;
- vi. Innovativeness level from web-scraping.

Preliminary results for the pilot service Health homecare

Results for the homecare services in three countries (Italy, the Netherlands and Portugal) are reported. In the case of the Netherlands, Table 17 shows that the service is not recorded for (and possibly not provided by) almost the half of Dutch health agencies and hospitals. Overall, the percentages of both entities across the innovativeness levels are coherent, especially for levels 0 and 2. The web-scraping exercise finds a higher number of highly innovative agencies (26 out of 147, that is 18 %) rather than hospitals (only 8 %). Also, this result does make sense, given the heterogeneity of hospitals, for which it is reasonable to expect few cases with dedicated services, described in a specific web page (thus identifiable via web-scraping). In addition, in the Netherlands — as also in the other countries of this pilot — health agencies are responsible for homecare services. It is then correct to find in their case a more balanced distribution of observations (23 %,

15 % and 18 % respectively) across innovativeness levels.

In order to check for the consistency of this first set of results, this table has been produced for two additional countries as well. In Table 18 results for Italy are displayed. For Italy, the distribution of observations across innovativeness levels is very different. There are lower number of units not providing homecare services (especially in hospitals) and a higher number of level 2 cases. For both health agencies and hospitals, almost half of the sample is identified at level 2, which are tele-medicine services.

Table 17: Baseline results for innovation in homecare in the Netherlands

| Innovativeness level | Health agencies | % of total | Hospitals | % of total |
|----------------------|-----------------|------------|-----------|------------|
| 0 | 65 | 44 % | 30 | 46 % |
| 1 | 34 | 23 % | 21 | 32 % |
| 2 | 22 | 15 % | 9 | 14 % |
| 3 | 26 | 18 % | 5 | 8 % |
| Total | 147 | 100 % | 65 | 100 % |

Source: StarPIN data.

Table 18: Baseline results for innovation in homecare in Italy

| Innovativeness level | Health agencies | % of total | Hospitals | % of total |
|----------------------|-----------------|------------|-----------|------------|
| 0 | 12 | 14 % | 15 | 32 % |
| 1 | 16 | 18 % | 3 | 6 % |
| 2 | 42 | 47 % | 22 | 46 % |
| 3 | 19 | 21 % | 7 | 16 % |
| Total | 89 | 100 % | 47 | 100 % |

Source: StarPIN data.

In Table 19, the summary results for Portugal are shown by including all the potential units of observation (hospitals, health agencies and other public institutions). In this case, missing observations — that is health units for which no innovativeness level is found — are almost absent and also the lowest level of innovativeness (traditional homecare services) is found only twice.

Table 19: Baseline results for innovation in homecare in Portugal

| Innovativeness level | Units of observation in Portugal | % of total |
|----------------------|----------------------------------|------------|
| 0 | 3 | 7 % |
| 1 | 2 | 4 % |
| 2 | 18 | 40 % |
| 3 | 22 | 49 % |
| Total | 45 | 100 % |

Source: StarPIN data.

Concerning Italian health agencies (ASL) without homecare services, a validation can be found in the statistics of Rapporto Osservasalute 2016, based on Ministerial data, reporting around 90 % of ASL providing homecare services. However, these official data do not report on the 'quality' of homecare services. The only measure concerns the share of population getting access to the service, at the NUTS2 regional level. Thus, correlation could be expected between our higher levels of innovativeness and a higher share of population with homecare services, based on the idea that more innovative services are more likely to get to a wider share of population.

A check was then performed on 26 units in Italy featuring a level 3 of innovativeness (19 ASL and 7 hospitals) with a special attention to the regions where they are located (Table 20).

Table 20: Results for Italy aggregated at the national and NUTS2 level

| NUTS2 region | Number of level 3 units | Share of regional population potentially reached by homecare services (year 2014) |
|-----------------------|-------------------------|---|
| Italian average | | 58 % |
| Sicily | 7 | 62 % |
| Veneto | 4 | 66 % |
| Lombardy | 3 | 60 % |
| Lazio | 3 | 25 % |
| Liguria | 2 | 99 % |
| Piedmont | 1 | 88 % |
| Friuli Venezia Giulia | 1 | 43 % |
| Valle d'Aosta | 1 | 10 % |
| Emilia Romagna | 1 | 100 % |
| Marche | 1 | 41 % |
| Molise | 1 | 74 % |
| Calabria | 1 | 10 % |

Source: StarPIN data.

This administrative information can hardly be directly associated with our innovativeness levels. Indeed, the high number of units getting to level 3 in Sicily suggests performing further qualitative checks. Indeed, the Sicily region is found to have developed several projects on tele-medicine and other distance health services. However, the existence of projects on homecare services might not imply the existence of an established service for the entire population. In other words, high numbers for that Italian region are not to be considered a mistake, but rather the identification of an innovative service more potential than actual.

On the contrary, Dutch units with highest level of innovativeness for homecare appear to be more distributed on the territory. For instance, the five hospitals getting to level 3 are located in four different provinces: Overijssel, Utrecht, North Holland and two in Gelderland.

In both the Italian and the Dutch cases, however, a main issue is represented by the very high number of units with level 0. Being that the keywords for the identification of level 1 are very generic (e.g. cure, aid, monitoring, domiciliary), it is quite unlikely that so many cases are excluded in the web-scraping. While for health agencies, as already recalled, this might be due to a number of units not providing the service, for hospitals there might be a technical issue. In fact, 30 Dutch hospitals and 15 Italian hospitals are not found when looking for these basic keywords. A qualitative check has allowed to find that for some of these hospitals (also very important, like <http://www.mauriziano.it> in Turin, Italy, and <http://www.havenpolikliniek.nl> in Rotterdam, the Netherlands) the web-scraping stopped after checking very few pages and did not find/assign correctly the level of innovativeness. For instance, the hospital in Turin has at least two web pages including the word 'domiciliare' (home) that could not be scraped.

Conclusions

This chapter has illustrated the potential that the StarPIN methodology has in terms of data generation with implications for our knowledge and understanding of innovation in public sector innovation. The main findings are described and commented on in detail, stemming from the pilot application of the methodology described in Chapter 6, to five EU countries, with illustrative analyses based on the newly collected data. The bulk of data collection and elaborations refers to the Waste management service domain, built and refined in two pilot countries in particular, Italy and the Netherlands, and extended to Portugal, Slovakia and Sweden. The pilot studies differ in terms of

their coverage of municipalities in the examined countries. However, it is shown that the methodology has the potential to capture at least part of the innovation phenomenon in this sector. Some expectations, based on previous surveys and data collections at a more aggregate level, have been largely confirmed, as in the case of the positive performance of Nordic countries. Nevertheless, the data allow the description of a more substantial heterogeneity across countries in terms of innovativeness of service provision. Furthermore, the StarPIN methodology allows exploring innovation at a much more detailed and significant level, as it allows some inference on innovativeness of municipalities in the provision of individual services. Of course data can and should be aggregated using appropriate upscaling procedures, some of which are discussed in this chapter. However, the richness of micro-level data can be exploited and, once aggregated at the regional and national level, data allow comparing the development and provision of specific services. A simple aggregation exercise carried out in section 7.2 of this chapter permits the production of a clear ranking of the five countries in terms of innovativeness of waste management services, and shows how regions are distributed around national means for each individual country. Although with a lesser depth, the methodology has been applied also to another service domain, Health homecare, and results are shown with reference to Italy, Netherlands and Portugal. The validation process and the cross-sectional analysis of the newly collected data together with administrative data and statistical data which are connected to public service innovation in general and to innovation in the two pilot services in particular, are discussed in view of future applications and research work.

8

Learning from the application of the methodology

By taking into consideration the activities described so far, as well as the challenges and obstacles met in the process of developing and applying the StarPIN methodology, this chapter focuses on the improvements to be introduced and the limitations to be dealt with in future data-collection efforts and related research.

Main difficulties

The main difficulties to be faced had to do with the following domains:

- i. *Service specificities*: the distinctive technical and organisational features of each service translate into complex and specific terminology and jargon, making it difficult to identify the appropriate keywords to describe the characteristics of the service and its delivery systems. Further complexities arise from the heterogeneity of website structures which in turn affects the ability of researchers/data miners to capture information on the service.
- ii. *Institutional characteristics*: the more or less hierarchical institutional setting and distribution of tasks and responsibilities across administrative levels; the size and organisation of the (public and private) entities providing the service.
- iii. *Context specific characteristics and policies*: the tradition/style, effort, and quality of communication (the awareness for transparency/accountability). As the method is based on the description/illustration of the service provided, it might be biased by differences in the communication; strategy/tradition/style/institutional settings that are at least partially country and region specific.
- iv. *Accessibility of statistical and administrative data*. This may require an in-depth knowledge of the local language and of statistical and institutional structure involved in data collection and provision practices. The main difficulty in this respect is to identify and to involve local experts in the national definition of the steps for the methodology application.

In order to extend the application of the methodology to a larger number of countries, one needs a 'tailor-made' translation of keywords and an appropriate selection of URLs for each specific national, regional and even local context. There are margins for improvement, e.g. to decrease the number of cases in which the service description was not effective, and led to no detection of a given service characteristic (the 0 level). The importance of such difficulties is confirmed when comparing the web-scraping results and the manual search results. Additional data collection difficulties arise when websites are hard to explore or not adequately functioning.

A well-organised list of keywords appeared to be of paramount importance to significantly reduce the number of both false positive and negatives. Apparently minor errors in the way keywords are ordered and separated one from the other, as well as the syntax used to write them down, may

impede the functioning of web-scraping procedures ⁽³³⁾.

Once the web-scraping procedure starts, the execution time cannot be estimated in a univocal way, as it depends not only on the number of URLs and keywords inserted in the scripts, but also on the structure of the websites and on technical limitations of available programs ⁽³⁴⁾.

Data collection was particularly difficult in the case of Italy and Portugal. Websites used by municipalities in these countries appear to be characterised by more complex structures. For this reason, scraping took longer in these cases. By contrast, exploring webpages of Swedish municipalities has not raised any problems and took less time. Websites exhibited a simpler structure that allowed the program to work in a shorter time without any particular hitches.

Lessons learnt through the testing activity

In applying the methodology to the two pilot services — Municipal waste collection and Health homecare — the StarPIN team identified several critical activities and routines that needed to be undertaken to solve specific problems. Further improvements to the application of the method and to the method structure itself will surely emerge while extending its application to other public services and to all the European countries.

One of the critical activities required for preparing the launch of a web-scraping for capturing the level of innovativeness of a public service, as described and foreseen by the methodology, is an accurate analysis of the wording used to describe the service and to accompany the user/beneficiary along the choices for activating the service delivery. A mother tongue, national/statistical/sectoral expert is highly recommended in the setting up of the web-scraping keyword list and URL list for each service/country.

Some practical aspects require attention:

- i. Know the conditions to use the tools — languages, programs, scripts, etc.;
- ii. Access to websites, orthography of URLs;
- iii. Time, machine time, a devoted machine with no stand by;
- iv. How many websites to be visited in one go, how to avoid 'visit' duplication;
- v. Network speed:
 - e.g. with a PC and fast internet connection, exploring each website requires up to one hour and a half
 - with a portion of machine and network speed, 15 to 20 websites can be explored in 1 hour
- vi. Updating a Logbook to keep track of the experience and codify knowledge accumulated through learning by doing.

⁽³³⁾ All keywords must be included in quotation marks and separated from the operators “:”, “1” and “,”. If the order of the operators is not respected the program will signal an error, preventing its execution. The same applies for the list of URLs. They must be written in the form “http://www” using the quotation marks and each URL must be separated with a comma. Also in this case, the syntax of the text is necessary to allow the reading, otherwise an error will appear on the program interface.

⁽³⁴⁾ The analysis of websites proceeds in two steps at different website levels, firstly identifying the pages containing the keywords to be submitted to subsequent scrutiny in the second step. During the process if the time required to explore a given page exceeds 8 000 000 ms the program will continue skipping the page concerned. For this reason, scraping may not be possible for those municipalities with more complex website structure during the first step of the execution, or some subdomains are not included in the results file because, once the first step is executed, the second is not executable. In some circumstances, the change of the URLs from “http” to “https” caused a system crash.

To decrease the number of missing cases:

- double check the URL name and where necessary enter manually the website in order to avoid/bypass navigation limitations (e.g. in some cases https has to be substituted with http in order to be allowed to access the website; in other circumstances, the first page or menu URL has a suffix which requires an additional ‘click’ to bring the visitor into the website);
- control the unexpected interruptions (checking the correspondence of the last subject result obtained with the last subject URL of the list) on which to act manually.

To create the conditions to replicate the web scraping, the following issues need to be taken into consideration:

- i. Software scripts: Ready-made computer codes should be made available to the user, with clear instructions on how to run them;
- ii. User manual: Detailed guidelines on how to collect and process the data;
- iii. Logbook: Illustrating step by step how problems have been solved;
- iv. Plan of time and resources to be employed to evaluate sustainability of the approach;
- v. Setting up an off-line Helpdesk based on a group of multidisciplinary researchers, to provide information needed in the web-scraping process. Resources to be employed to set up a help desk might not be urgently required for potential extension of the pilot phase of the methodology, but have to be planned and found as the service and country application expands.

Directions for future improvements

In this section, a selection of potential directions for future development of the methodology are mentioned or roughly sketched.

(a) Further decrease the number of URLs for which the web scraping does not succeed:

Possible reasons for a **missing answer** from a website:

- i. The service is not delivered;
- ii. The website is under maintenance;
- iii. The crawler fails to retrieve the contents of the URL for other reasons.

A first obvious explanation is that the URL itself is no longer valid, for instance because the organisation has changed the domain name of its website. In the particular case of public sector organisations, this is a rare case since domain names are usually directly linked to the actual name of the organisations⁽³⁵⁾. Problems might occur due to a restructuring of a public administration (e.g. the merger of several municipalities). An obvious but hard-to-pursue remedy is to always keep the list of URLs up to date. Missing URLs should be checked manually. Note that in practice old websites will usually automatically be redirected to new websites. A crawler will also follow the redirect hence will still be able to identify the new website via the old one if it is using a redirect http status code in its headers. When a site would use a link with explanatory text on their site, it will be interpreted as a normal link.

Websites might also be **temporarily** unavailable due to technical problems. Most web crawling suites store non-functioning URLs and have a loop to automatically reiterate those URLs at a later stage. At the level of individual webpages (hence not on the aggregate level of websites) error ‘404’

⁽³⁵⁾ For instance, <https://www.comune.roma.it> is the official website for the municipality of Rome, <https://www.amsterdam.nl> for the municipality of Amsterdam.

occur more often. This is particularly the case for relatively large websites (> 20 000 webpages) that have a lot of internal references. Sometimes, internal links are no longer functioning. The web-crawler will then not find the specific destination page.

Websites could also be **structurally** unavailable. This might either be due to the policy of the webmaster with regard to web-crawlers or to technical reasons. In the first case, a webmaster might instruct robots (such as web-crawlers) to not crawl parts of the website, or not crawl the website altogether. This assumes that the robot reveals its identity (via robots.txt) but it is good web practice to do so (and to explicitly mention the name of the party that is collecting the data). In our experience public sector organisations rarely block access for robots. In the private sector, price comparison sites are a known issue. Intranets — the parts of an organisation network that is only accessible to the internal staff — are private by definition. Hence, these are the ‘hidden’ parts in the access procedures. Web-crawlers are not supposed to have access to intranets, and in the unlikely case that access is technically feasible, access (due to faulty security settings) is not allowed from a legal point of view. For the particular use of public services — services that are offered to the public at large, hence are public by definition — the presence of the ‘hidden’ part of the network from the public organisation is just not relevant.

Technical issues might occur either because websites use very old or very new web technology. For example, a known problem for web-crawlers is the use of the outdated Table-element rather than cascading style sheets (CSS). Such websites will not be rendered correctly by modern web browsers hence will also not be parsed in the right way (or not be parsed at all) by web-crawlers. In the case of the Netherlands, no single case was found of missing websites of public organisations due to outdated technology (in fact there were no instances of missing URLs altogether). In the case of Italy some cases have been indeed found. These were all very outdated static websites from small municipalities. As a general rule it would be safe to consider the provisioning of online public services at such websites at the lowest level in the innovation classification ⁽³⁶⁾.

The other way around, crawlers might also have limitations when it comes to websites that use very new web technologies. For example, the increasing use of JavaScript (e.g. CSS and HTML in nowadays being injected into JavaScript rather than the other way around) has greatly challenged the automatic testing and crawling of websites. However, the use of headless browsers has largely solved this problem, and the recent introduction of the Puppeteer tool (which was also used in the pilots) has further simplified the use and stability of web crawling scripts. Hence, the lesson is that web crawlers should always use the state-of-the-art web technology. Obviously, this is a moving target.

As shown, the set of circumstances impeding the obtaining of a ‘response’ from a website (no service available, websites under maintenance or out of order, non-efficient websites or web-crawlers) is quite extensive and variegated. Non-technical observers can hardly distinguish each circumstance one from another. Hence, the only solution that could be undertaken is to randomly test missing answer cases with manual web-scraping and check the actual frequency of each individual case.

(b) Focus on technological levels of websites as innovativeness indicators

A partially complementary pattern to improve the outcome of web-scraping, along with the enhancement of techniques to reduce failures to explore the **contents** of websites is to rely on information on **technological properties** of websites. Two objects of analysis can be addressed here:

b.1) The server technology level, monitored via the upload/download speed, the access time and other indicators (see D4 for details on data available from Google Analytics on this aspect). Data collection on the technology performance for the websites visited can be conducted at a much finer level with web-scraping techniques.

⁽³⁶⁾ It is technically feasible to deal with the issue of framesets by writing a piece of tailor-made code that follows each frame in the frameset, identifies the mainframe (this is where the content is) and then navigates through the website while treating each frame as the separate webpage.

b.2) Assessment of how technology level can affect service provision, e.g. the ability to interact with citizens/beneficiaries.

Future analysis and measurement could be envisaged to collect these data to complement content-based analyses through web-scraping.

From a technology point of view, the critical bottleneck with regard to capacity is not on the side of the web crawler. To crawl a sizeable number of websites, several servers are needed but capacity could also be acquired in a cost-efficient way over the internet (i.e. in the cloud, however see hereafter). The bottleneck is rather on the side of the web servers that run the websites that are being crawled. The issue is that too-frequent requests on the web server will stress server load too much. This could even bring the server to a grinding halt and as a consequence the website will become unresponsive to other visitors. This is much akin to a denial of service attack. To avoid such problems (i.e. to scrape the web 'in a polite way') delays should be built in between the sequential requests from a crawler. Such delays can be set as static values (e.g. 3s) but since websites vary drastically in the number of requests they can handle it is much more effective (and efficient) to dynamically adjust the delay according to the current web server load, based on the latency from the previous request.

While applying such dynamically adjusted latencies, it is certainly possible to crawl all Dutch municipal websites **completely** (n = 350) in 1 month. This covering the entire content of the website (hence up to the lowest level) and including machine-readable formats such as Word and pdf files but excluding graphics and audio and video files. Depending on the frequency of changes in the content of the websites, subsequent runs could be done much faster (up to 75 %) because the previous content is being cached, thus only the altered content would have to be crawled again. Note that the throughput time for Dutch municipalities is relatively high due to the particular circumstance that a lot of municipal websites are currently being hosted on one web server. As a consequence, the delay times have to be set quite high. If each website is being hosted on a dedicated server, crawling can be done in parallel. The nearly 8 000 municipal websites in Italy could be crawled in 1-2 months if this condition applies.

Obviously, parallel crawling would require the deployment of a large-scale IT infrastructure that is dedicated to web crawling. In theory, such capacity could be temporarily hired in the Cloud. However the availability of a sophisticated IT infrastructure is not sufficient to be able to run large-scale web crawling projects. The bottleneck is rather in the expertise that is needed. It requires specialist knowledge to run and constantly tweak the performance of such web crawling campaigns. There are specialised firms that offer these services. In order to optimise the performance of the system (especially to minimise runtimes) they generally use their own servers rather than (shared) cloud services.

On the other hand, a much simpler solution was adopted in the pilots on the basis that it can be run locally on one computer and that does not require specialised knowledge. To minimise run time a targeted script was used that only crawls what is relevant to a particular topic (i.e. a particular public service) and processes the target data on the fly (i.e. it counts the frequency of a predefined set of keywords on each selected webpage). Thus a website is only partially crawled and the actual content is not being downloaded. This brings down the runtime to about 2 days for 100 websites (1 day for subsequent runs) and the results are readily available for further analysis.

(c) Potential application of machine learning for ex post definition of keywords and detecting new innovation levels

The StarPIN web-scraping methodology has so far drawn on a priori knowledge on specific public service domains in specific national/territorial contexts. It did have to rely on local experts to select relevant organisations (i.e. URLs) and to define a list of keywords and link those keywords (or rather: combinations of keywords) to service levels.

Instead of making use of experts another trajectory would be to automate the classification process. Obviously, this approach heavily relies on IT and requires in-depth knowledge of data science. The

main advantage over the expert approach is that the actual classification is done after the data has been acquired — all data is available offline. This makes it in principle a more flexible approach. The classification can be readjusted without having to collect the data (i.e. crawl the websites) again. However, this assumes that the entire content from the websites is being crawled and stored ('site mirroring') (See Chapter 2 for a discussion of advantages and limitations of this methodological option). Moreover, the (large amount of) raw data also needs to be cleaned and parsed before it can be analysed. In the case of headless browsers, which basically emulate the behaviour of a human agent who visits a website, the information that is being shown to a visitor can be easily separated from the rest of the content. In the case of 'site mirroring' the information that is shown to the visitor (the actual text on the webpage) has to be filtered out afterwards.

Once the data have been cleaned up and parsed, various techniques can be used to automatically classify the data. In terms of machine learning, supervised learning algorithms can be implemented⁽³⁷⁾. It is supervised because the algorithm is fed with training data which defines which data is relevant and which is not. In this particular case the training data consists of a set of webpages that is known to describe the public service at hand, and a set of webpages that is known **not** to describe the same public service.

The compilation of a proper set of training data is a critical task (it basically determines the outcome of the machine learning exercise). How much training data is needed depends on the complexity of the problem and the type of algorithm that is used. As a basic rule of thumb, several hundred of positive and negative cases would be needed. The more sophisticated the algorithm, the more training data is needed. Non-linear algorithms for instance (e.g. random forest, artificial neural networks) could need 10 times for data (hence thousands instead of hundreds of cases). Hence, if the problem at hand refers to a relatively small amount of cases (e.g. less than 500) the set of training data would already approach the size of the total number of cases and it would make little sense to automatically classify the remaining cases.

The actual automated classification process advances in two steps:

1. The identification and selection of 'relevant' entities from the target data. Here: relevant keywords from a webpage.
2. The calculation of the propensity scores on one or more predefined classes. This calculation depends on the distribution of the relevant entities and the relationships (e.g. 'distance') between those entities. In the ideal case, webpages are assigned 100 % to one specific class (one service level for one particular public service). In practice, webpages will often be assigned to multiple classes and a threshold value will be used to reassign the webpage to one particular class. Even distributions constitute a residual category of 'undetermined cases' which must still be classified by hand. Obviously, if this residual category is too large the added value of the use of automated classification over manual classification is limited.

For both steps (1) and (2) a (different) wide range of techniques and algorithms are available (hence many more combinations). The selection of the most suitable techniques and the optimal combination thereof, is a second critical task in the deployment of machine learning. Ironically, in the presence of the numerous sophisticated algorithms, this is ultimately an empirical inquiry, that is, a matter of trial and error.

(d) Possible application of web-scraping to user benefit search

The use of scraping and/or automated classification is not limited to webpages. It might, as well, be applied to other types of online data, such as social media. Social media are interesting data source because they cover the demand side (the users) rather than the supply side (the public services). Machine learning applied to communications on social media (e.g. tweets on Twitter, posts on Facebook) have been used to map public opinion on particular topics. The reference here is

⁽³⁷⁾ For a detailed overview of supervised (classification) and unsupervised (clustering) machine learning algorithms see: te Velde, R. A. (2017), 'Enterprise profiling. Using data analytics to classify firms based on innovative activities', Utrecht: Dialogic, Working paper written for Eurostat (project ref. ESTAT/G/2015/006).

specifically to ‘opinion mining’ (or ‘sentiment analysis’). With sentiment analysis, a large number of conversations are scanned to automatically determine the underlying emotion that is present.

The usual disclaimers apply. Large sets of good training data are required and, given the huge challenge to minimise the set of indeterminate cases — which is potentially very large in the case of opinion mining, actually a lot of training data is needed. Moreover, the complexity of social expressions (e.g. use of negative terms in a positive sense, use of irony and sarcasm) requires the use of sophisticated algorithms — hence ever more training data.

In the particular case of (the assessment of the performance of) public services the potential use of sentiment analysis is most probably low. This is because sentiment analysis works best in cases of **temporary peaks of conflicting opinions**. However, the assessment of public services is probably much more even in nature, both over time and in terms of the distribution of opinions (no extremes). The total amount of statements about social services on social media will be limited anyway.

(e) Relying on open data as a service and as a means to access service contents

Governments increasingly publish their data via open data portals. Data on public services is usually also present in the open datasets. One advantage of this type of data is that it is generally well-structured and provided with metadata. Open data on public services could be useful as a baseline to assess the validity of the outcome of web-based measurements.

However, coverage of open data of public services has so far been very limited. For instance, in the UK — which has been one of the frontrunners in the open data movement and which has one of the largest repositories of open data in the world — out of the 45 479 available datasets only 406 refer to waste, and within this subset only 131 refer to towns and cities (0.27 % on a total of 48 000 towns and cities). For the Netherlands and Italy comparable results were found, respectively 86 and 207 hits on ‘waste’ on a total of 12 891 and 20 814 open datasets. Similarly to the UK, in the Netherlands only nine datasets refer to municipalities (0.23 % of 388 municipalities).

(f) Additional potential developments:

The use of web-based measurements and the launch of traditional measurement procedures (e.g. surveys) do not exclude each other. There is, in fact, much to be gained in the combination of the two types of measurements. Given the superior performance of web-based measurements in terms of scaling, one obvious setup would be to use surveys to validate (parts of) the results of the web-based measurements. Such a combined operating process could be largely automated. Results from the web-based measurements could be directly forwarded (e.g. by means of an online survey) to contact persons from the organisation at hand. The eventual errors in the results of the web-based measurements could be corrected. At a more general level, the feedback of the respondents could be used to adjust the list of keywords and/or the scoring of the service levels.

Conclusions

This concluding chapter described the challenges and obstacles met in the process of developing and applying the StarPIN methodology. Potential improvements and possible limitations to be dealt with in data-collection efforts and in future research have been discussed. The practical application experience has also led to identify a set of procedures and routines that need to be put into action to correct and/or prevent errors, and to improve the efficiency of the data collection process. The acknowledgement of problems and solutions encountered in the application of the methodology is a fundamental part of a learning process that led to an improvement our own approach. There is ample room for ameliorating data-collection and our understanding of public sector innovation. This is, of course, subject matter for future research and applications of this methodology.

References

- Ahmed, S. A. and Ali, S. M. (2006), 'People as partners: Facilitating people's participation in public-private partnerships for solid waste management', *Habitat International* 30:781-796.
- Arora, S. K., Youtie, J., Shapira, P. and Li, Y. (2014), Does 'Pivoting' Matter for Innovation? The Usefulness of Strategic Business Changes in Analyzing Growth in Small and Medium-Sized Green Goods Firms, Annual Conference of the International Association for Management of Technology. Washington DC: IAMT.
- Arundel, A., Bloch, C. and Ferguson, B. (2016), Methodologies for measuring innovation in the public sector, OECD Blue Sky Forum on Science and Innovation Indicators, 19-21 September 2016, Ghent, Belgium, pp. 1-22.
- Bakhshi, H. and Mateos-Garcia, J. (2016, 11 February), Tech Nation 2016: much more than disruption, available at <http://www.nesta.org.uk/blog/tech-nation-2016-findings>.
- Bakhshi, H., Davies, J. and Mateos-Garcia, J. (2015), *The Net Effect: Using social media data to understand the impact of a conference on social networks*, London: Nesta.
- Barras, R. (1986), 'Towards a theory of innovation in services', *Research Policy* 15, 161-173.
- Barras, R. (1990), 'Interactive innovation in financial and business services: the vanguard of the service revolution', *Research Policy* 19, 215-237.
- Bason, C. (2010), *Leading public sector innovation. Co-creating for a better society*, Bristol, England: Policy Press.
- Bason, C. (2017), *Leading Public Design: (Discovering) Human-Centered Governance*, Paperback, 15 April 2017.
- Bean, C. (2015), *Independent review of UK economic statistics: interim report*, London: Chancellor of the Exchequer.
- Beaudry, C., Héroux-Vaillancourt, M. and Rietsch, C. (2016), Validation of a web mining technique to measure innovation in high technology Canadian industries, Ghent: OECD Blue Sky Forum 2016.
- Bel, G., Fageda, X. and Mur, M. (2013), 'Why Do Municipalities Cooperate to Provide Local Public Services? An Empirical Analysis', *Local Government Studies* 39(3):435-454.
- Bianchi, A., Cozza, C., Marin, G., Rienstra, Y., te Velde, R. A. and Zanfei, A. (2017), 'Review of the literature on public service innovation', StarPIN working paper D1.
- Bianchi, A., Cozza, C., Koppers, W., Marin, G., Perani, G., te Velde, R. A., Zanfei, A. and Zecca, E. (2018), 'A new instrument for data collection on Public Innovation in EU countries', StarPIN working paper D2.
- Bianchi, A., Cozza, C., Jelacic, N., Koppers, W., Marin, G., Perani, G., te Velde, R. A., Zanfei, A. and Zecca, E. (2018), 'Report on the methodology testing activity in a selected group of EU countries', StarPIN working paper D3.
- Bianchi, A., Perani, G. and Zanfei, A. (2018), 'Measuring ICT innovation in the public sector and its

- integration with Eurostat's activities', StarPIN working paper D4.
- Bianchi, A., Marin, G. and Zanfei A. (2018), 'New perspectives in public service innovation', in Gallouj, F. and Djellal, F. (eds) (2018), *A Research Agenda for Service Innovation*, E. Elgar Publishers.
- Blaug, R., Horner, L., Kenyon, A. and Lekhi, R. (2006), *Public value and local communities. A literature review*, The Work Foundation.
- Blei, D. M., Ng, A. Y. and Jordan, M. I. (January 2003), Lafferty, J., (ed.) 'Latent Dirichlet Allocation', *Journal of Machine Learning Research* 3 (4-5): pp. 993-1022, doi:10.1162/jmlr.2003.3.4-5.993.
- Botta, E. and Kozluk, T. (2014), 'Measuring Environmental Policy Stringency in OECD Countries: A Composite Index Approach', OECD Economics Department Working Papers, No 1177, OECD Publishing, Paris.
- Bovens, M. and Zouridis, S. (2002), 'From street-level to system-level bureaucracies', *Public Administration Review* 62(2), 174-184.
- Brennenraedts, R. and te Velde, R. A. (2012), *Internet as data source. Feasibility Study on Statistical Methods on Internet as a Source of Data Gathering*, Luxembourg: Publications Office of the European Union.
- Brennenraedts, R., Holland, C., Batenburg, R., den Hertog, P., te Velde, R. A., Jansen, S. and Brinkkemper, S. (2008), *Go with the dataflow! Analysing the Internet as a data source (IaD)*, Main report, The Hague: Ministry of Economic Affairs.
- Buccioli, A., Montinari, N. and Piovesan, M. (2015), 'Do Not Trash the Incentive! Monetary Incentives and Waste Sorting', *Scandinavian Journal of Economics* 117(4):1204-1229.
- Bugge, M., Mortensen, P. and Bloch, C. (2011), Measuring Public Innovation in Nordic Countries, MEPIN, NIFU, Norway, <http://www.nifu.no/files/2012/11/NIFUrapport2011-40.pdf>.
- Bulkley, H., Watson, M., Hudson, R. and Weaver, P. (2005), 'Governing municipal waste: towards a new analytical framework', *Journal of Environmental Policy & Planning* 7(1):1-23.
- Bulkley, H., Watson, M. and Hudson, R. (2007), 'Modes of governing municipal waste', *Environment and Planning A* 39:2733-2753.
- Capgemini (2016), *eGovernment benchmark 2016. Turning point for eGovernment development in Europe?*, Brussels: European Commission.
- Chesbrough, H. W. (2003), *Open Innovation: The new imperative for creating and profiting from technology*, Boston, Harvard Business School Press.
- Chesbrough, H. (2011), *Open Services Innovation: Rethinking Your Business to Grow and Compete in a New Era*, Harvard Business School Press, Boston.
- Cleary, J. (2009), 'Life cycle assessments of municipal solid waste management systems: A comparative analysis of selected peer-reviewed literature', *Environment International* 35:1256-1266.
- COFOG UN Classification of the Functions of Government, [https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Classification_of_the_functions_of_government_\(COFOG\)](https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Classification_of_the_functions_of_government_(COFOG)).
- Coombs, R. and Miles, I. (2000), 'Innovation, Measurement and Services: The New Problematique', in: Metcalfe, J. S. and Miles, I., (eds) *Innovation Systems in the Service Economy. Economics of Science, Technology and Innovation*, vol. 18. Springer, Boston, MA.
- CPA Statistical classification of products by activity, [https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_products_by_activity_\(CPA\)](https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_products_by_activity_(CPA)).
- Dataprovider (2017), Hearbeat (Dataprovider), available on 24 May 2017 at <https://www.dataprovider.com>: <https://www.dataprovider.com/industries/business-information>
- Devi, T. and Aparna, K. (2016), 'A Survey on Web Mining: Overview, Techniques, Tools, and Applications', *International Journal for Research in Applied Science & Engineering Technology*, 4(1),

257-264.

De Vries, E. J. (2006), 'Innovation in services in networks of organisations and in the distribution of services', *Research Policy*, vol. 35, n. 7, pp 1037-1051.

De Vries, H., Bekkers, V. and Tummers, L. (2016), 'Innovation in the public sector: a systematic review and future research agenda', *Public Administration* Vol. 94, No. 1, 2016 (146-166).

Diplaris, S., Papadopoulos, S. and Gkatziaki, V. (2017), *The Wikirate project. Scalable Analytics Techniques for User Contributions*, Thessaloniki: CERTH.

Djellal, F., Gallouj, F. and Miles, I. (2013), 'Two decades of research on innovation in services: Which place for public services?', *Structural Change and Economic Dynamics* 27 (2013), 98-117.

Droegeet, H., Heras Forcada, M. A. and Hildebrand, D. (2009), 'Innovation in Services: Present Findings, and Future Pathways', *Journal of Service Management* 20(2):131-155.

Eaton, K. (2012, 15 March), How One Second Could Cost Amazon \$1.6 Billion In Sales, available on 25 May 2017, at <https://www.fastcompany.com>: <https://www.fastcompany.com/1825005/how-one-second-could-cost-amazon-16-billion-sales>.

Edison, H., bin Ali, N. and Torkar, R. (2013), 'Towards innovation measurement in the software industry', *The Journal of Systems and Software*, vol. 86, issue 5.

Edquist, C. and Hommen, L. (2000), *Public technology procurement and innovation theory*, Kluwer Academic Publishers, New York.

EEA (2015), *Waste prevention in Europe — the status in 2014*, EEA Report 6/2015, European Environment Agency, Copenhagen, Denmark.

EEA (2013), *Municipal Waste management*, by Ferraris, M. and Paleari, S., CERIS-CNR, <https://www.eea.europa.eu/publications/managing-municipal-solid-waste/italy-municipal-waste-management/view>.

EIONET European Topic Centre on Air Pollution and Climate Change Mitigation—EEA European Environment Agency (2017), Report, Italian section, <https://www.eea.europa.eu/countries-and-regions/italy>.

Elia, V., Gnoni, M. G. and Tornese, F. (2015), 'Designing Pay-As-You-Throw schemes in municipal waste management services: A holistic approach', *Waste Management* 44:188-195.

Elias, S. (2016), Value Confusion: The Problem of Lean in Public Services — Lean Competency System — Lean Thinking, <http://www.leancompetency.org/lcs-articles/value-confusion-problem-lean-public-services>, 1-14.

Ellinger, A. E., Lynch, D. F. and Hansen, J. D. (2003), 'Firm size, web site content, and financial performance in the transportation industry', *Industrial Marketing Management* 32, 177-185.

Enticott, G. and Walker, R. M. (2008), 'Sustainability, performance and organizational strategy: an empirical analysis of public organizations', *Business Strategy and the Environment* 17:79-92.

Ernst & Young (2017), *Public sector innovation. From ideas to actions*.

EUPHA (2013), PHIRE — *Public Health Innovation and Research in Europe*, Summary Report.

European Commission (2011), *Innobarometer 2010. Analytical Report Innovation in Public Administration*.

European Commission (2013), European Public Sector Innovation Scoreboard EPSIS.

European Commission (2017a), *eGovernment Benchmark 2017. Taking stock of user-centric design and delivery of digital public services in Europe*, DG Connect, Final background report — volume 2.

European Commission (2017b), *eGovernment Benchmark 2017. Taking stock of user-centric design and delivery of digital public services in Europe*, DG Connect, Final insight report — volume 1

European System of National and Regional Accounts ESA (2010),

<https://ec.europa.eu/eurostat/web/esa-2010>.

Eurostat's VIP ADMIN Project (2018), Analysis of the legal and institutional environment in the EU Member States and EFTA Countries, Work Package 1, https://ec.europa.eu/eurostat/cros/content/wp1-access-administrative-data_en.

Eurostat's VIP ADMIN Project (2018), Good practices in accessing, using and contributing to the management of administrative data, Work Package 1, https://ec.europa.eu/eurostat/cros/content/wp1-access-administrative-data_en.

Evison, T. and Read, A. D. (2001), 'Local Authority recycling and waste — awareness publicity/promotion', *Resources, Conservation and Recycling* 32:275-291.

Fuglsang, L. (2008), 'Capturing the benefits of open innovation in public innovation: a case study', *International Journal of Services Technology and Management* 9 (3-4), 234-248.

Fuglsang, L. (2010), 'Bricolage and invisible innovation in public service innovation', 2010/1, *Journal of Innovation Economics* 1(5), 67-87.

Fuglsang, L. and Rønning, R. (2015), 'On innovation patterns and value-tensions in public services', *The Service Industries Journal* 35 (2015), Issue 9, 467-482.

Gallouj, F. (1998), 'Innovating in reverse: services and the reverse product cycle', *European Journal of Innovation Management*, Vol. 1, Issue: 3, pp. 123-138.

Gallouj, F. and Weinstein, O. (1997), 'Innovation in services', *Research Policy*, 1997, vol. 26, issue 4-5, 537-556.

Gallouj, F. (2002), 'Interactional innovation: a neoschumpeterian model', in: Sundbo, J. and Fuglsang, L. (Eds.), *Innovation as Strategic Reflexivity*, Routledge, London, New York, pp. 29-56.

Gallouj, F. and Toivonen, M. (2011), 'Elaborating the characteristics-based approach to service innovation: Making the service process visible', *Journal of Innovation Economics* 2011/2, (8), pp. 33-58.

Gallouj, F. and Zanfei, A. (2013), 'Innovation in public services: Filling a gap in the literature', *Structural Change and Economic Dynamics*, vol. 27, n. 4, 2013, pp. 89-97.

Gascò, M. (2017), 'Living labs: Implementing open innovation in the public sector', *GIQ Government Information Quarterly*, Vol. 34, Issue 1, January 2017, 90-98.

Garsiel, T. and Irish, P. (2011, 5 August), How Browsers Work: Behind the scenes of modern web browsers, available on 24 May 2017, at: <https://www.html5rocks.com/en/tutorials/internals/howbrowserswork>.

Gault, F. (2012), 'User innovation and the market', *Science and Public Policy* 39, 118-128.

Giljum, S., Lieber, M. and Doranova, A. (2016), The Eco-Innovation Scoreboard: 2014 and 2015 versions, Technical note, available at: http://ec.europa.eu/environment/ecoap/scoreboard_en

GIQ (2017), *Special issue on open innovation in the public sector*, no. 34.

Gök, A., Waterworth, A. and Shapira, P. (2015), 'Use of web mining in studying innovation', *Scientometrics*, 653-671.

Hatry, H., Winnie, R. E., and Fisk D. M. (1981), *Practical Program Evaluation for State and Local Governments*, Urban Institute Press, Washington D.C.

Hatry, H., Wholey, J. and Newcomer, K. (eds.) (1994), *Handbook of Practical Program Evaluation*, Jossey Bass, San Francisco.

Hoekstra, R., ten Bosch, O. and Hartevelde, F. (2010), Automated Data Collection from Web Sources for Official Statistics: First Experiences, The Hague: CBS Dutch Statistics.

Howells, J. (2001), 'The nature of innovation in services', *Innovation and productivity in services*, OECD, 55-79.

- Hughes, A., Moore, K. and Kataria, N. (2011), *Innovation in Public Sector Organisations. A pilot survey for measuring innovation across the public sector*, NESTA, March 2011.
- ISPRA (2017), *Rapporto rifiuti urbani* (Report on solid waste), Istituto Superiore per la Protezione e la Ricerca Ambientale, Rome.
- ITU (2014), *Manual for Measuring ICT Access and Use by Households and Individuals*, International Telecommunication Union, Geneva.
- Jaffe, A. B., Newell, R. G. and Stavins, R. N. (2005), 'A tale of two market failures: Technology and environmental policy', *Ecological Economics* 54(2-3):164-174.
- Jager, C.-J. and Vankan, A. (2015), *Digitale vaardigheden op de arbeidsmarkt* (digital skills on the labour market), Utrecht: Dialogic.
- Jager, C.-J., Bongers, F., Veldkamp, J., te Velde, R. A., Vankan, A. and Veldman, J. (2014), *Marktonderzoek Mariënborg* (market research Mariënborg data science master course), Utrecht: Dialogic.
- Joseph, K. (2006), 'Stakeholder participation for sustainable waste management', *Habitat International* 30:863-871.
- Kankanhalli, A., Zuiderwijk, A. and Tayi, G. K. (2017), 'Open innovation in the public sector: A research agenda', *Government Information Quarterly*, Volume 34, Issue 1, January 2017, 84-89.
- Karagiannidisa, A., Xirogiannopoulou, A. and Tchobanoglous, G. (2008), 'Full cost accounting as a tool for the financial assessment of Pay-As-You-Throw schemes: A case study for the Panorama municipality, Greece', *Waste Management* 28(12):2801-2808.
- Katz, S. J. and Cothey, V. (2006), 'Web indicators for complex innovation systems', *Research Evaluation* 15(2), 85-95.
- Kemp, R. and Pearson, P. (2008), Final Report MEI project about measuring eco-innovation. Measuring Eco-Innovation FP6 Project, available at: http://cordis.europa.eu/docs/publications/1245/124548931-6_en.pdf
- Kiparsky, M., Thompson, B. H., Binz, C., Sedlak, D. L., Tummers, L. and Truffer, B. (2016), 'Barriers to innovation in urban wastewater utilities: attitudes of managers in California', *Environmental Management* 57:1204-1216.
- Kim, J. H. (2012), 'A Hyperlink and Semantic Network Analysis of the Triple Helix (University-Government-Government-Communication Structure of Nanotechnology)', *Journal of Computer-Mediated Communication*, 17, 152-170.
- Kloek, W. and Vâju, S. (2013), 'The use of administrative data in integrated statistics', Eurostat.
- Libaers, D., Hicks, D. and Portery, A. L. (2010), 'A taxonomy of small firm technology commercialization', *Industrial and Corporate Change*, 1-35. doi:doi:10.1093/icc/dtq039
- Lie, D. H., Hulstijn, J., op den Akker, R. and Nijholt, A. (1998), A Transformational Approach to Natural Language Understanding in Dialogue Systems, available on 25 May 2017, at https://www.researchgate.net/publication/2473679_A_Transformational_Approach_to_Natural_Language_Understanding_in_Dialogue_Systems.
- Lombrano, A. (2009), 'Cost efficiency in the management of solid urban waste', *Resource, Conservation and Recycling* 53:601-611.
- Loukis, E. et al. (2017), 'Promoting open innovation in the public sector through social media monitoring', *GIQ Government Information Quarterly*, Volume 34, Issue 1, January 2017, 99-109.
- Luckin, D. and Sharp, L. (2004), Remaking Local Governance through Community Participation? The Case of the UK Community Waste Sector, *Urban Studies* 41(8):1485-1505.
- Manning, C. D., Raghavan, P. and Schütze, H. (2008), *Introduction to Information Retrieval*, Cambridge University Press, Cambridge.

- Mateos-Garcia, J. (2016, 4 February), Arloesiadur: an innovation analytics experiment, available at <http://www.nesta.org.uk/blog/arloesiadur-innovation-analytics-experiment>
- Mateos-Garcia, J., Bakhshi, H. and Lenel, M. (2014), *A map of the UK games industry*, London: Nesta.
- Miles, I. (1998), *Services, Systems and Innovation — Cinderella comes to the ball*, Services and Systems Worksop, CRIC, Manchester, March.
- Misuraca, G., Codagnone, C. and Rossel, P. (2013), 'From Practice to Theory and back to Practice: Reflexivity in Measurement and Evaluation for Evidence-based Policy Making in the Information Society', *Government Information Quarterly*, 30.
- Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P. and Bhattacharjee, B. (2007), Measurement and analysis of online social networks. IMC '07 Proceedings of the 7th ACM SIGCOMM conference on Internet measurement (pp. 29-42), San Diego (CA): ACM.
- Misuraca, G. and Viscusi, G. (2015), 'Shaping public sector innovation theory: an interpretative framework for ICT-enabled governance innovation', *Electronic Commerce Research*, 15(3).
- Moore, M. H. (1995), *Creating Public Value. Strategic Management in Government*, Cambridge, MA: Harvard University Press.
- Moore, M. H. and Beninton, J. (2010), *Public value: Theory and practice*, London: Palgrave, Macmillan.
- Moore, M. H. (2012), Recognizing Public Value: Developing a Public Value Account and a Public Value Scorecard, <http://publiccommons.ca/public/uploads/literature/Moore-9.4.12.pdf>.
- Moore, M. H. (2013), *Recognizing Public Value*, Cambridge, MA: Harvard University Press.
- Mulgan, G. (2014), *Innovation in the public sector. How can public organisations better create, improve and adapt?*, NESTA, November 2014.
- NACE Statistical classification of economic activities in the European Community, [https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_economic_activities_in_the_European_Community_\(NACE\)](https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_economic_activities_in_the_European_Community_(NACE)).
- Nambisan, P. and Nambisan, S. (2009), 'Models of consumer value cocreation in health care', *Health care management review*, 34(4), 344-354.
- Nathan, M. and Rosso, A. (2012), *Measuring the UK's Digital Economy with Big Data*, London: Growth Intelligence/National Institute of Economic and Social Research.
- Neely, A. (2008), 'Exploring the Financial Consequences of the Servitization of Manufacturing', *Operations Management Research*, 1, 103-118.
- NESTA — Hughes et al. (2011), NESTA pilot survey of health organisations and local government organisations in the UK.
- Norden Nordic Council of Ministers (2010), *Health Innovation in the Nordic Countries*, Public Private Collaboration.
- Norden Nordic Council of Ministers (2011), *Innovative Public Procurement and Health Care*, Nordic Lighthouse Project.
- OECD (2005), *E-Government for Better Government*, Paris: OECD.
- OECD (2005), *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data*, third ed. OECD, Paris.
- OECD (2008), *Environmental Policy, Technological Innovation and Patents*, OECD Publishing, Paris.
- OECD (2011), *Public Research Institutions: Mapping Sector Trends*, Paris.
- OECD (2013), *Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy*,

Paris: OECD.

OECD (2014), *Measuring the Digital Economy: A New Perspective*, Paris: OECD.

Oslo Manual (2018), *Guidelines for Collecting, Reporting and Using Data on Innovation*, 4th Edition, The Measurement of Scientific, Technological and Innovation Activities, Paris:OECD.

Ohemeng, F. K. and Grant, J. K. (2008), *When markets fail to deliver: an examination of the privatization and de-privatization of water and wastewater services delivery in Hamilton, Canada*. Canadian Public Administration/Administration Publique du Canada 51(3):475-499.

Oostrom, L., Walker, A. N., Staats, B., Slootbeek-Van Laar, M., Ortega Azurduy, S. and Rooijakkers, B. (2016), *Measuring the internet economy in The Netherlands: a big data analysis*, The Hague: CBS Statistics Netherlands.

Osborne, S. P. (Ed.) (2010), *The New Public Governance? Emerging Perspectives on the Theory and Practise of Public Governance*, Routledge, London.

Osborne, S. P. and Brown, L. (Eds.) (2013), *Handbook of Innovation and Change in Public Services*, Edward Elgar, Cheltenham.

Osborne, S. P., Radnor, Z. and Nasi, G. (2013), 'A New Theory for Public Service Management? Toward a (Public) Service-Dominant Approach', *American Review of Public Administration* 43(2) 135-158.

Page, L. and Brin, S. (1998), 'The Anatomy of a Large-Scale Hypertextual Web Search Engine', *Computer Networks*, 30(1-7), 107-111.

Pine, B. J. and Gilmore, J. H. (1999), 'The Experience Economy', *Harvard Business Review* 76(6):176, 97-105.

Reggi, L., Arduini, D., Biagetti, M. and Zanfei, A. (2013), 'How Advanced are Italian Regions in Terms of Public e-Services? The Construction of a Composite Indicator to Analyze Patterns of Innovation in the Public Sector', *Telecommunications Policy*, vol.38, 2014, pp. 514-529.

Reddick, C. G., Takeoka Chatfield, A. and Ojo, A. (2017), 'A social media text analytics framework for double-loop learning for citizen-centric public services: A case study of a local government', *G/IQ Government Information Quarterly*, Volume 34, Issue 1, January 2017, 110-125.

Rhodes, R. A. W. (1996), 'The New Governance Governing without Government', *Political Studies*, 44, 652-667.

Rousta, K., Bolton, K., Lunding, M. and Dahlén, L. (2015), 'Quantitative assessment of distance to collection point and improved sorting information on source separation of household waste', *Waste Management* 40:22-30.

Savona, M. and Steinmueller, W. E. (2013), 'Service output, innovation and productivity: A time-based conceptual framework', *Structural Change and Economic Dynamics*, vol. 27, pp. 118-132.

Schut, M., te Velde, R. A. and Flipse, R. (2017), *Monitoren van de implementation van de Omgevingswet op basis van raadsverslagen (monitoring the implementation of the Environmental Licencing Act based on town council minutes)*, Rotterdam: Notubiz.

Scupola, A. and Zanfei, A. (2016), 'Governance and innovation in public sector services: The case of the digital library', *Government Information Quarterly*, vol. 33, Issue 2, April 2016, pp. 237-249, DOI 10.1016/j.giq.2016.04.005

Seyfang, G. and Smith, A. (2007), 'Grassroots innovations for sustainable development: Towards a new research and policy agenda', *Environmental Politics* 16:(4):584-603.

Seri, P. and Zanfei, A. (2013), 'The co-evolution of ICT, skills and organization in public administrations: Evidence from new European country-level data', *Structural Change and Economic Dynamics*, Elsevier, vol. 27(C), 160-176.

ServPPIN (2012), *The Contribution of Public and Private Services to European Growth and Welfare*,

- and the Role of Public-Private Innovation Networks, FP7, See <http://www.servppin.com>.
- Shapira, P., Gök, A. and Salehi, F. (2016), Graphene enterprise: mapping innovation and business development in a strategic emerging technology, *J. Nanopart Res.*, 18(269).
- Shin, S. K. and Sanders, G. L. (2006), 'Denormalization strategies for data retrieval from data warehouses', *Decision Support Systems*, 42(1):267-282, October 2006.
- Skelcher, C., Mathur, N. and Smith, M. (2005), 'The Public Governance of Collaborative Spaces', *Public Administration* 83(3):573-596.
- SNA System of National Accounts (2008), Ch. F. The general government sector and its subsectors.
- Sørensen, E. and Torfing, J. (2012), 'Introduction Collaborative Innovation in the Public Sector', *The Innovation Journal: The Public Sector Innovation Journal*, vol. 17(1), 1-14.
- Souders, S. (2007), *High Performance Websites*, Farnham: O'Reilly Media.
- Speight, V. L. (2015), Innovation in the water industry: barriers and opportunities for US and UK utilities, *WIREs Water* 2:301-313.
- Stahel, W. R. (1997), The service economy: 'Wealth without resource consumption'?, *Philosophical Transactions of the Royal Society A: Mathematical Physical and Engineering Sciences* 355(1728):1309-1319.
- Stanford Medicine (2017), *Health Trends Report. Harnessing the Power of Data in Health*.
- Statistics Canada, <http://www.statcan.gc.ca/pub/12-539-x/2009001/administrative-administratives-eng.htm>.
- Stewart Loane, S., Webster, C. M. and D'Alessandro, S. (2015), Identifying consumer value co-created through social support within online health communities, *Journal of Macromarketing*, 35(3), 353-367.
- T&C (2012), *Trends and Challenges in Public Sector Innovation in Europe*, Thematic Report 2012 under Specific Contract for the Integration of INNO Policy TrendChart with ERAWATCH (2011-2012), by Leon, R. et al.
- Tauchmann, H., Clausen, H. and Oelmann, M. (2009), 'Do organizational forms matter? Innovation and liberalization in the German wastewater sector', *Journal of Policy Modeling* 31:863-876.
- te Velde, R. A. and Lelie, T. (2017a), *The economic footprint of Booking.com in The Netherlands*. Amsterdam: Booking.com.
- te Velde, R. A. (2017b), *Enterprise profiling. Using data analytics to classify firms based on innovative activities*, Utrecht: Dialogic, Working paper written for Eurostat (project ref. ESTAT/G/2015/006).
- The Conference Board (2006), *The Conference Board Help Wanted OnLine® (HWOL)*, available on May 24, 2017, at <https://www.conference-board.org/data/helpwantedonline.cfm>.
- UN (2014), *Manual for measuring e-government, United Nations Economic Commission for Africa and Partnership on measuring ICT for development*, Addis Ababa.
- van der Lei, T. E. and Cunningham, S. W. (2006), Use of the internet for future-oriented technology analysis. Second International Seville Seminar on Future-Oriented Technology Analysis, Seville: IPTS.
- Vandermerwe, S. and Rada, J. (1988), 'Servitization of business: Adding value by adding services', *European Management Journal*, vol. 6, Issue 4, Winter 1988, pp. 314-324.
- Vankan, A., Janssen, M. and Veldman, J. (2014), *Clusteranalyse Rotterdam foodsector*, Utrecht: Dialogic.
- Vargo S. L. and Lusch, R. F. (2006), 'Service-dominant logic: What it is, what it is not, what it might be', in Vargo, S. L. and Lusch, R. F. (eds.), *The Service Dominant Logic of Marketing: Dialog*,

- Debate, and Directions*, Publisher: M. E. Sharpe, Inc., pp. 43-55.
- Vargo, S. L. and Lusch, R. F. (2008), 'Service Dominant Logic: Continuing the evolution', *J. of the Acad. Mark. Sci.* (2008) 36:1-10.
- Veefkind, V., Hurtado-Albir, J., Angelucci, S., Karachalios, K. and Thumm, N. (2012), 'A new EPO classification scheme for climate change mitigation technologies', *World Patent Information* 34:106-111.
- Victor, J., Montgomery, A. H. and Lubell, M. (2016), *The Oxford Handbook of Political Networks*, Oxford: Oxford University Press.
- WHO (2016), *Global diffusion of eHealth: making universal Health coverage achievable*, Report of the third global survey on eHealth.
- Windrum, P. and Garcia Goni, M. (2008), 'A Neo-Schumpeterian Model of Health Services Innovation', *Research Policy* 37(4):649-672.
- World Health Organisation (2010), *Telemedicine: opportunities and developments in Member States*, Report on the second global survey on eHealth 2009.
- World Health Organisation (2016), *Global diffusion of eHealth: making universal health coverage achievable*, Report of the third global survey on eHealth.
- Xu, R., Sun., Q. and Si, W. (2015), 'The Third Wave of Public Administration: The New Public Governance', *Canadian Social Science* Vol. 11, No. 7, 2015, pp. 11-21.
- Yin, L. A., Arora, S., Youtie, J. and Shapiro, P. (2016), 'Using web mining to explore Triple Helix influences on growth in small and mid-size firms', *Technovation*, doi:<http://dx.doi.org/10.1016/j.technovation.2016.01.002>
- Youtie, J., Hicks, D., Shapira, P. and Horsely, T. (2012), 'Pathways from Discovery to Commercialization: Using Web Sources to Track Small and Medium-sized Enterprise Strategies in Emerging Nanotechnologies', *Technology Analysis & Strategic Management*, 24(10).
- Zuurmond, A. (2005), *Organisational Transformation Through the Internet*, Cambridge: Cambridge University Press.

Annex 1: Classification Grid

| COFOG | COFOG definition | CPA | CPA definition |
|--------|---|----------|---|
| 01.1.1 | Executive and legislative organs | 84.11.11 | Executive and legislative services |
| | | 84.11.19 | Other general (overall) public services |
| | | 84.11.29 | Other supporting services for the government |
| 01.1.2 | Financial and fiscal affairs | 84.11.12 | Financial and fiscal services |
| | | 84.11.19 | Other general (overall) public services |
| | | 84.11.29 | Other supporting services for the government |
| 01.1.3 | External affairs | 84.21.11 | Administrative external affairs-related services, diplomatic and consular services abroad |
| | | 84.21.13 | Foreign military aid-related services |
| 01.2 | Foreign economic aid | 84.21.12 | Foreign economic aid-related services |
| 01.3.1 | General personnel services | 84.11.21 | General personnel services for the government |
| 01.3.2 | Overall planning and statistical services | 84.11.13 | Overall economic and social planning and statistical services |
| | | 84.11.19 | Other general (overall) public services |
| | | 84.11.29 | Other supporting services for the government |
| | | 71.12.20 | Project management services for construction projects |
| | | 84.11.19 | Other general (overall) public services |
| 01.3.3 | Other general services | 84.11.29 | Other supporting services for the government |
| | | 41.00.10 | Residential buildings |
| | | 41.00.20 | Non-residential buildings |
| | | 41.00.30 | Construction works for residential buildings (new works, additions, alterations and renovation works) |
| | | 41.00.40 | Construction works for non-residential buildings (new works, additions, alterations and renovation works) |
| | | 68.20.11 | Rental and operating services of own or leased residential real estate |
| | | 68.20.12 | Rental and operating services of own or leased non-residential real estate |
| | | 68.31.11 | Residential buildings and associated land sale services on a fee or contract basis, except of time-share ownership properties |
| | | 68.31.12 | Time-share properties sale services on a fee or contract basis |
| | | 68.31.13 | Residential vacant land sale services on a fee or contract basis |
| | | 68.31.14 | Non-residential buildings and associated land sale services on a fee or contract basis |
| | | 68.31.15 | Non-residential vacant land sale services on a fee or contract basis |
| | | 68.31.16 | Real estate appraisal services on a fee or contract basis |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|------------------|----------|--|
| | | 68.32.11 | Residential property management services on a fee or contract basis, except of time-share ownership properties |
| | | 68.32.12 | Time-share property management services on a fee or contract basis |
| | | 68.32.13 | Non-residential property management services on a fee or contract basis |
| | | 91.01.11 | Library services |
| | | 91.01.12 | Archive services |
| 01.4 * | Basic research | 84.11.14 | Government services to fundamental research |
| | | 72.11.11 | Research and experimental development services in health biotechnology |
| | | 72.11.12 | Research and experimental development services in environmental and industrial biotechnology |
| | | 72.11.13 | Research and experimental development services in agricultural biotechnology |
| | | 72.11.20 | Research and development originals in biotechnology |
| | | 72.19.11 | Research and experimental development services in mathematics |
| | | 72.19.12 | Research and experimental development services in computer and information sciences |
| | | 72.19.13 | Research and experimental development services in physical sciences |
| | | 72.19.14 | Research and experimental development services in chemistry |
| | | 72.19.15 | Research and experimental development services in earth and related environmental sciences |
| | | 72.19.16 | Research and experimental development services in biological sciences |
| | | 72.19.19 | Research and experimental development services in other natural sciences |
| | | 72.19.21 | Research and experimental development services in nanotechnology |
| | | 72.19.29 | Other research and experimental development services in engineering and technology, except biotechnology |
| | | 72.19.30 | Research and experimental development services in medical sciences |
| | | 72.19.40 | Research and experimental development services in agricultural sciences |
| | | 72.19.50 | Research and development originals in natural sciences and engineering, except for biotechnology |
| | | 72.20.11 | Research and experimental development services in economics and business |
| | | 72.20.12 | Research and experimental development services in psychology |
| | | 72.20.13 | Research and experimental development services in law |
| | | 72.20.19 | Research and experimental development services in other social sciences |
| | | 72.20.21 | Research and experimental development services in languages and literature |
| | | 72.20.29 | Other research and experimental development services in humanities |
| | | 72.20.30 | Research and development originals in social sciences and humanities |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|--|----------|---|
| 01.4 * | Administration of R & D policies and related funds, intended to increase personal well-being, related to basic research | 84.11.14 | Government services to fundamental research |
| | | 84.12.11 | Administrative educational services |
| | | 84.12.12 | Administrative health care services |
| | | 84.12.13 | Administrative housing and community amenity services |
| | | 84.12.14 | Administrative recreational, cultural and religious services |
| 01.4 * | Administration of R & D policies and related funds, intended to improve economic performance, related to basic research | 84.13.11 | Administrative agriculture-, forestry-, fishing- and hunting-related services |
| | | 84.13.12 | Administrative fuel- and energy-related services |
| | | 84.13.13 | Administrative mining- and mineral resources-, manufacturing- and construction-related services |
| | | 84.13.14 | Administrative transport- and communications-related services |
| | | 84.13.15 | Administrative services related to the distributive and catering trades, hotels and restaurants |
| | | 84.13.16 | Administrative services related to tourism affairs |
| | | 84.13.17 | Administrative multipurpose development project services |
| | | 84.13.18 | General administrative economic, commercial and labour affairs-related services |
| 01.5 * | R & D General public services | 84.11.14 | Government services to fundamental research |
| | | 71.20.12 | Testing and analysis services of physical properties |
| | | 71.20.13 | Testing and analysis services of integrated mechanical and electrical systems |
| 01.5 * | Administration of R & D policies and related funds, intended to increase personal well-being, related to general public services | 84.11.14 | Government services to fundamental research |
| 01.5 * | Administration of R & D policies and related funds, intended to improve economic performance, related to general public services | 84.11.14 | Government services to fundamental research |
| 01.6 | General public services n.e.c. | 84.11.13 | Overall economic and social planning and statistical services |
| | | 84.11.19 | Other general (overall) public services |
| | | 84.11.29 | Other supporting services for the government |
| 01.7 | Public debt transactions | 84.11.12 | Financial and fiscal services |
| | | 84.11.13 | Overall economic and social planning and statistical services |
| | | 84.11.19 | Other general (overall) public services |
| | | 84.11.29 | Other supporting services for the government |
| 01.8 | Transfer of a general character between different levels of government | 84.11.12 | Financial and fiscal services |
| | | 84.11.13 | Overall economic and social planning and statistical services |
| | | 84.11.19 | Other general (overall) public services |
| | | 84.11.29 | Other supporting services for the government |

| COFOG | COFOG definition | CPA | CPA definition |
|----------|--|----------|---|
| 02.1 | Military defence | 84.22.11 | Military defence services |
| 02.2 | Civil defence | 84.22.12 | Civil defence services |
| 02.3 | Foreign military aid | 84.21.13 | Foreign military aid-related services |
| 02.4 * | R & D defence (except administration of defence-related R & D policies and related funds) | 72 | Scientific research and development services |
| 02.4 * | Administration of defence-related R & D policies and related funds | 84.22.1 | Defence services |
| 02.5 | Defence n.e.c. | 84.22.1 | Defence services |
| 03.1 | Police services | 84.24.11 | Police services |
| | | 84.24.19 | Other public order and safety affairs-related services |
| | | 71.20.19 | Other technical testing and analysis services |
| 03.2 | Fire-protection services | 84.25.11 | Fire-fighting and fire-prevention services |
| | | 84.25.19 | Other fire brigade services |
| | | 71.20.19 | Other technical testing and analysis services |
| 03.3 | Law courts | 84.23.11 | Law courts-related administrative services |
| | | 71.20.19 | Other technical testing and analysis services |
| 03.4 | Prisons | 84.23.12 | Administrative services related to detention or rehabilitation of criminals |
| | | 71.20.19 | Other technical testing and analysis services |
| 03.5 | R & D Public order and safety | 72 | Scientific research and development services |
| 03.6 | Public order and safety n.e.c. | 84.24.1 | Public order and safety services |
| 04.1.1 * | Operation of weather forecasting and geodesic surveys | 74.90.14 | Weather forecasting and meteorological services |
| | | 74.90.13 | Environmental consulting services |
| | | 71.11.31 | Urban planning services |
| | | 71.11.32 | Rural land planning services |
| | | 71.11.33 | Project site master planning services |
| | | 71.11.41 | Landscape architectural services |
| | | 71.11.42 | Landscape architectural advisory services |
| | | 71.12.11 | Engineering advisory services |
| | | 71.12.31 | Geological and geophysical consulting services |
| | | 71.12.32 | Geophysical services |
| | | 71.12.33 | Mineral exploration and evaluation services |
| | | 71.12.34 | Surface surveying services |
| | | 71.12.35 | Map-making services |
| | | 09.90.19 | Support services to other mining and quarrying n.e.c. |
| 04.1.1 * | General economic and commercial affairs (except foreign commercial affairs, operation of weather forecasting and geodesic surveys) | 84.13.18 | General administrative economic, commercial and labour affairs-related services |
| 04.1.1 * | General foreign commercial affairs | 84.21.11 | Administrative external affairs-related services, diplomatic and consular services abroad |
| 04.1.2 | General labour affairs | 84.13.18 | General administrative economic, commercial and labour affairs-related services |
| 04.2.1 | Agriculture | 84.13.11 | Administrative agriculture-, forestry-, fishing- and hunting-related services |
| 04.2.2 * | Forest fire fighting and fire prevention services | 02.40.10 | Support services to forestry |

| COFOG | COFOG definition | CPA | CPA definition |
|----------|--|----------|---|
| 04.2.2 * | Forestry (except forest fire fighting and fire prevention services) | 84.13.11 | Administrative agriculture-, forestry-, fishing- and hunting-related services |
| | | 02.40.10 | Support services to forestry |
| 04.2.3 * | Operation of fish hatcheries | 84.13.11 | Administrative agriculture-, forestry-, fishing- and hunting-related services |
| | | 03.00.72 | Support services to aquaculture |
| 04.2.3 * | Fishing and hunting (except operation of fish hatcheries) | 84.13.11 | Administrative agriculture-, forestry-, fishing- and hunting-related services |
| 04.3.1 | Coal and other solid mineral fuels | 84.13.12 | Administrative fuel- and energy-related services |
| 04.3.2 | Petroleum and natural gas | 84.13.12 | Administrative fuel- and energy-related services |
| 04.3.3 | Nuclear fuel | 84.13.12 | Administrative fuel- and energy-related services |
| 04.3.4 | Other fuels | 84.13.12 | Administrative fuel- and energy-related services |
| 04.3.5 * | Operation of non-enterprise-type electricity supply systems | 35.11.10 | Electricity |
| | | 35.12.10 | Transmission services of electricity |
| | | 35.13.10 | Distribution services of electricity |
| | | 35.14.10 | Trade services of electricity |
| 04.3.5 * | Electricity (except operation of non-enterprise-type electricity supply systems) | 84.13.12 | Administrative fuel- and energy-related services |
| | | 42.22.11 | Long-distance electricity power lines and communication lines |
| | | 42.22.12 | Local electricity power lines and communication lines |
| | | 42.22.13 | Power plants |
| | | 42.22.21 | Construction works for long-distance electricity power lines and communication lines |
| | | 42.22.22 | Construction works for local electricity power lines and communication lines |
| | | 42.22.23 | Construction works for power plants |
| | | 71.12.13 | Engineering services for power projects |
| 04.3.6 * | Operation of non-enterprise-type non-electricity supply systems | 35.30.11 | Steam and hot water |
| | | 35.30.12 | Steam and hot water supply services through mains |
| | | 35.30.21 | Ice, including ice for cooling (i.e. non-food) purposes |
| | | 35.30.22 | Cooled air and chilled water supply services |
| 04.3.6 * | Non-electric energy (except operation of non-enterprise-type non-electricity supply systems) | 84.13.12 | Administrative fuel- and energy-related services |
| 04.4 | Mining, manufacturing and construction | 84.13.13 | Administrative mining- and mineral resources-, manufacturing- and construction-related services |
| 04.5.1 * | Operation of non-enterprise-type road transport systems | 49.31.21 | Urban and suburban scheduled road transport services of passengers |
| | | 49.31.22 | Mixed mode urban and suburban scheduled transport services of passengers |
| | | 49.39.11 | Interurban scheduled road transport services of passengers |
| | | 49.39.12 | Interurban special-purpose scheduled road transport services of passengers |
| | | 49.39.13 | Other special-purpose scheduled road transport services of passengers |
| | | 49.39.33 | Non-scheduled local bus and coach charter services |

| COFOG | COFOG definition | CPA | CPA definition |
|----------|--|----------|--|
| | | 49.39.34 | Non-scheduled long distance bus and coach charter services |
| | | 49.39.39 | Passenger land transport services n.e.c. |
| 04.5.1 * | Operation of non-enterprise-type road transport facilities | 33.17.19 | Repair and maintenance services of other transport equipment n.e.c. |
| | | 52.21.21 | Bus station services |
| | | 52.21.22 | Highway operation services |
| | | 52.21.23 | Bridges and tunnel operation services |
| | | 52.21.24 | Parking lot services |
| | | 52.21.25 | Towing services for private and commercial vehicles |
| | | 52.21.29 | Other services incidental to road transportation |
| 04.5.1 * | Road transport (except operation of non-enterprise-type road transport systems and facilities) | 84.13.14 | Administrative transport- and communications-related services |
| | | 71.20.14 | Technical inspection services of road transport vehicles |
| 04.5.2 * | Operation of non-enterprise-type water transport systems and facilities | 50.10.11 | Sea and coastal passenger water transport services by ferries |
| | | 50.10.12 | Sea and coastal passenger water transport services on cruise ships |
| | | 50.10.19 | Other sea and coastal passenger water transport services |
| | | 50.20.22 | Towing and pushing services on sea and coastal waters |
| | | 50.40.22 | Towing and pushing services on inland waters |
| | | 50.30.11 | Inland passenger water transport services by ferries |
| | | 50.30.19 | Other inland passenger water transport services |
| 04.5.2 * | Supporting activities to water transport | 52.22.11 | Port and waterway operation services (except cargo handling) on sea and coastal waters |
| | | 52.22.12 | Inland waterway operation services (except cargo handling) |
| | | 52.22.13 | Pilotage and berthing services on sea and coastal waters |
| | | 52.22.14 | Pilotage and berthing services on inland waters |
| | | 52.22.15 | Vessel salvage and refloating services on sea and coastal waters |
| | | 52.22.16 | Vessel salvage and refloating services in inland waters |
| | | 52.22.19 | Other services incidental to water transportation |
| | | 33.15.10 | Repair and maintenance services of ships and boats |
| 04.5.2 * | Water transport (except supporting activities to water transport and except operation of non-enterprise-type water transport systems and facilities) | 84.13.14 | Administrative transport- and communications-related services |
| 04.5.3 * | Operation of non-enterprise-type railway transport systems | 33.17.11 | Repair and maintenance services of railway locomotives and rolling-stock |
| | | 49.10.19 | Other passenger rail transport services, interurban |
| | | 49.20.19 | Other railway transport services of freight |
| | | 49.31.10 | Urban and suburban railway transport services of passengers |
| | | 52.21.19 | Other services incidental to railway transportation |
| 04.5.3 * | Operation of non-enterprise-type railway transport facilities | 33.17.11 | Repair and maintenance services of railway locomotives and rolling-stock |
| | | 33.17.19 | Repair and maintenance services of other transport equipment n.e.c. |
| | | 52.21.11 | Railway pushing or towing services |

| COFOG | COFOG definition | CPA | CPA definition | | |
|----------|---|----------|--|----------|--|
| 04.5.3 * | Railway transport (except operation of non-enterprise-type railway transport systems and facilities) | 52.21.19 | Other services incidental to railway transportation | | |
| | | 84.13.14 | Administrative transport- and communications-related services | | |
| | | 42.12.10 | Railways and underground railways | | |
| | | 42.12.20 | Construction works for railways and underground railways | | |
| | | 42.13.10 | Bridges and tunnels | | |
| | | 42.13.20 | Construction works for bridges and tunnels | | |
| 04.5.4 * | Operation of non-enterprise-type air transport systems and facilities | 51.10.11 | Domestic scheduled air transport services of passengers | | |
| | | 51.10.12 | Domestic non-scheduled air transport services of passengers, except for sightseeing | | |
| | | 51.10.13 | International scheduled air transport services of passengers | | |
| | | 51.10.14 | International non-scheduled air transport services of passengers | | |
| | | 51.21.11 | Scheduled air transport services of intermodal containers | | |
| | | 51.21.12 | Air transport services of letters and parcels | | |
| | | 51.21.13 | Scheduled air transport services of other freight | | |
| | | 51.22.12 | Space transport services of freight | | |
| | | 04.5.4 * | Support activities to air transport | 33.16.10 | Repair and maintenance services of aircraft and spacecraft |
| | | | | 52.23.11 | Airport operation services, excluding cargo handling |
| 52.23.12 | Air traffic control services | | | | |
| 52.23.19 | Other service activities incidental to air transportation | | | | |
| 52.23.20 | Services incidental to space transportation | | | | |
| 04.5.4 * | Air transport (except support activities to air transport and except operation of non-enterprise-type air transport systems and facilities) | 84.13.14 | Administrative transport- and communications-related services | | |
| 04.5.5 * | Operation of non-enterprise-type pipeline and other transport systems and facilities | 49.50.11 | Transport services via pipeline of crude or refined petroleum and petroleum products | | |
| | | 49.50.12 | Transport services via pipeline of natural gas | | |
| | | 49.50.19 | Transport services via pipeline of other goods | | |
| 04.5.5 * | Pipeline and other transport (except operation of non-enterprise-type pipeline and other transport systems and facilities) | 84.13.14 | Administrative transport- and communications-related services | | |
| 04.6 | Communication | 52.21.30 | Services incidental to transportation via pipelines | | |
| | | 84.13.14 | Administrative transport- and communications-related services | | |
| | | 53.10.11 | Postal services under universal service obligation related to newspapers and periodicals | | |
| | | 53.10.12 | Postal services under universal service obligation related to letters | | |
| | | 53.10.13 | Postal services under universal service obligation related to parcels | | |
| | | 53.10.14 | Post office counter services | | |
| | | 53.10.19 | Other postal services under universal service obligation | | |

| COFOG | COFOG definition | CPA | CPA definition |
|----------|--|----------|---|
| | | 53.20.19 | Other postal and courier services n.e.c. |
| | | 49.20.15 | Railway transport services of letters and parcels |
| | | 49.41.18 | Road transport services of letters and parcels |
| | | 61.90.10 | Other telecommunications services |
| | | 71.12.18 | Engineering services for telecommunications and broadcasting projects |
| 04.7.1 | Distributive trades, storage and warehousing | 84.13.15 | Administrative services related to the distributive and catering trades, hotels and restaurants |
| 04.7.2 | Hotels and restaurants | 84.13.15 | Administrative services related to the distributive and catering trades, hotels and restaurants |
| 04.7.3 * | Tourist offices | 79.90.11 | Tourism promotion services |
| | | 79.90.12 | Visitor information services |
| 04.7.3 * | Tourism | 84.13.16 | Administrative services related to tourism affairs |
| | | 79.90.20 | Tourist guide services |
| | | 79.90.32 | Reservation services for convention centres, congress centres and exhibit halls |
| | | 79.90.39 | Reservation services for event tickets, entertainment and recreational services and other reservation services n.e.c. |
| | | 79.90.31 | Time-share exchange services |
| 04.7.4 | Multi-purpose development projects | 84.13.17 | Administrative multipurpose development project services |
| 04.8 | R & D Economic affairs | 84.11.14 | Government services to fundamental research |
| 04.9 | Economic affairs n.e.c. | 84.11.13 | Overall economic and social planning and statistical services |
| | | 84.13.18 | General administrative economic, commercial and labour affairs-related services |
| 05.1 * | Administration of waste collection, treatment and disposal systems | 71.12.15 | Engineering services for waste management projects (hazardous and non-hazardous) |
| | | 71.20.11 | Composition and purity testing and analysis services |
| | | 74.90.13 | Environmental consulting services |
| 05.1 * | Waste management | 37.00.11 | Sewage removal and treatment services |
| | | 37.00.12 | Treatment services of cesspools and septic tanks |
| | | 37.00.20 | Sewage sludge |
| | | 38.11.11 | Collection services of non-hazardous recyclable waste, municipal |
| | | 38.11.19 | Collection services of non-hazardous recyclable waste, other |
| | | 38.11.21 | Collection services of non-hazardous non-recyclable waste, municipal |
| | | 38.11.29 | Collection services of non-hazardous non-recyclable waste, other |
| | | 38.11.31 | Non-recyclable non-hazardous municipal waste |
| | | 38.11.39 | Other non-recyclable non-hazardous waste |
| | | 38.11.41 | Vessels and other floating structures, for breaking up |
| | | 38.11.49 | Wrecks, other than vessels and floating structures, for dismantling |
| | | 38.11.51 | Glass waste |
| | | 38.11.52 | Paper and paperboard waste |
| | | 38.11.53 | Used pneumatic tyres of rubber |
| | | 38.11.54 | Other rubber waste |
| | | 38.11.55 | Plastic waste |
| | | 38.11.56 | Textile waste |
| | | 38.11.57 | Leather waste |
| | | 38.11.58 | Non-hazardous metal waste |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|---|----------|---|
| | | 38.11.59 | Other non-hazardous recyclable waste, n.e.c. |
| | | 38.11.61 | Services of transfer facilities for non-hazardous recyclable waste |
| | | 38.11.69 | Services of transfer facilities for other non-hazardous waste |
| | | 38.12.11 | Collection services of hazardous medical and other biohazardous waste |
| | | 38.12.12 | Collection services of other hazardous industrial waste |
| | | 38.12.13 | Collection services of hazardous municipal waste |
| | | 38.12.21 | Spent (irradiated) fuel elements (cartridges) of nuclear reactors |
| | | 38.12.22 | Pharmaceutical waste |
| | | 38.12.23 | Other medical hazardous waste |
| | | 38.12.24 | Hazardous chemical waste |
| | | 38.12.25 | Waste oils |
| | | 38.12.26 | Hazardous metal waste |
| | | 38.12.27 | Waste and scrap of primary cells, primary batteries and electric accumulators |
| | | 38.12.29 | Other hazardous waste |
| | | 38.12.30 | Services of transfer facilities for hazardous waste |
| | | 38.21.10 | Non-hazardous waste treatment for final disposal services |
| | | 38.21.21 | Sanitary landfill services |
| | | 38.21.22 | Other landfill services |
| | | 38.21.23 | Incineration services of non-hazardous waste |
| | | 38.21.29 | Other non-hazardous waste disposal services |
| | | 38.21.30 | Waste organic solvents |
| | | 38.21.40 | Ashes and residues from waste incineration |
| | | 38.22.11 | Nuclear waste treatment services |
| | | 38.22.19 | Other hazardous waste treatment services |
| | | 38.22.21 | Nuclear waste disposal services |
| | | 38.22.29 | Other hazardous waste disposal services |
| | | 81.29.11 | Disinfecting and exterminating services |
| | | 81.29.13 | Other sanitation services |
| | | 81.29.19 | Other cleaning services n.e.c. |
| | | 81.30.10 | Landscape services |
| | | 39.00.11 | Remediation and clean-up services, soil and groundwater |
| | | 39.00.12 | Remediation and clean-up services, surface water |
| | | 39.00.13 | Remediation and clean-up services, air |
| | | 39.00.14 | Building remediation services |
| | | 39.00.21 | Site remediation containment, control and monitoring services and other site remediation services |
| | | 39.00.22 | Other remediation services |
| | | 39.00.23 | Other specialised pollution control services |
| 05.2 * | Administration of waste water treatment systems | 71.12.15 | Engineering services for waste management projects (hazardous and non-hazardous) |
| | | 71.20.11 | Composition and purity testing and analysis services |
| | | 74.90.13 | Environmental consulting services |
| 05.2 * | Waste water management | 84.12.13 | Administrative housing and community amenity services |
| | | 81.29.11 | Disinfecting and exterminating services |
| | | 81.29.12 | Sweeping and snow removal services |
| | | 81.29.13 | Other sanitation services |
| | | 81.29.19 | Other cleaning services n.e.c. |
| | | 81.30.10 | Landscape services |
| | | 37.00.11 | Sewage removal and treatment services |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|---|----------|---|
| | | 37.00.12 | Treatment services of cesspools and septic tanks |
| | | 37.00.20 | Sewage sludge |
| | | 39.00.11 | Remediation and clean-up services, soil and groundwater |
| | | 39.00.12 | Remediation and clean-up services, surface water |
| 05.3 | Pollution abatement | 84.12.13 | Administrative housing and community amenity services |
| | | 81.29.13 | Other sanitation services |
| | | 81.29.19 | Other cleaning services n.e.c. |
| | | 39.00.21 | Site remediation containment, control and monitoring services and other site remediation services |
| | | 39.00.22 | Other remediation services |
| | | 39.00.23 | Other specialised pollution control services |
| 05.4 * | Protection of biodiversity and landscape (except operation of natural parks and reserves) | 91.04.11 | Botanical and zoological garden services |
| | | 81.30.10 | Landscape services |
| | | 71.11.41 | Landscape architectural services |
| 05.4 * | Operation of natural parks and reserves | 91.04.12 | Nature reserves services, including wildlife preservation services |
| 05.5 | R & D environmental protection | 72.11.12 | Research and experimental development services in environmental and industrial biotechnology |
| | | 72.19.15 | Research and experimental development services in earth and related environmental sciences |
| 05.6 | Environmental protection n.e.c. | 84.12.13 | Administrative housing and community amenity services |
| | | 71.11.31 | Urban planning services |
| | | 71.11.32 | Rural land planning services |
| | | 71.11.33 | Project site master planning services |
| | | 71.11.41 | Landscape architectural services |
| | | 71.11.42 | Landscape architectural advisory services |
| 06.1 * | Housing development (construction) | 41.00.10 | Residential buildings |
| | | 41.00.30 | Construction works for residential buildings (new works, additions, alterations and renovation works) |
| 06.1 * | Housing development (except construction) | 84.12.13 | Administrative housing and community amenity services |
| | | 71.11.10 | Plans and drawings for architectural purposes |
| | | 71.11.21 | Architectural services for residential building projects |
| | | 71.11.22 | Architectural services for non-residential building projects |
| | | 71.11.23 | Historical restoration architectural services |
| | | 71.11.24 | Architectural advisory services |
| | | 71.12.12 | Engineering services for building projects |
| | | 71.12.19 | Engineering services for other projects |
| | | 81.30.10 | Landscape services |
| 06.2 | Community development | 84.12.13 | Administrative housing and community amenity services |
| 06.3 * | Water supply | 35.30.11 | Steam and hot water |
| | | 35.30.12 | Steam and hot water supply services through mains |
| | | 35.30.21 | Ice, including ice for cooling (i.e. non-food) purposes |
| | | 35.30.22 | Cooled air and chilled water supply services |
| | | 36.00.11 | Drinking water |
| | | 36.00.12 | Non-drinking water |
| | | 36.00.20 | Treatment and distribution services of water through mains |
| | | 36.00.30 | Trade services of water through mains |
| 06.3 * | Construction of non-enterprise type water supply systems | 42.21.11 | Long-distance pipelines for fluids |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|--|----------|--|
| | | 42.21.12 | Local pipelines for fluids |
| | | 42.21.13 | Irrigation systems (canals); water main and line constructions; water treatment plants, sewage disposal plants and pumping stations |
| | | 42.21.21 | Construction works for long-distance pipelines |
| | | 42.21.22 | Construction works for local pipelines, including ancillary works |
| | | 42.21.23 | Construction works for irrigation systems (canals), water mains and lines, water treatment plants, sewage disposal plants and pumping stations |
| | | 42.21.24 | Water well drilling and septic system installation works |
| | | 42.91.10 | Coastal and port constructions, dams, locks and related hydro-mechanical structures |
| | | 42.91.20 | Construction works for coastal and port constructions, dams, locks and related hydro-mechanical structures |
| | | 42.99.19 | Other civil engineering constructions n.e.c. |
| | | 42.99.29 | Construction works for civil engineering constructions n.e.c. |
| | | 71.12.16 | Engineering services for water, sewerage and drainage projects |
| 06.3 * | Administration of water supply affairs | 84.12.13 | Administrative housing and community amenity services |
| 06.4 * | Installation of street lighting | 71.11.41 | Landscape architectural services |
| | | 71.12.14 | Engineering services for transportation projects |
| | | 71.12.17 | Engineering services for industrial and manufacturing projects |
| | | 43.21.10 | Electrical installation works |
| | | 42.11.10 | Motorways, roads, streets and other vehicular or pedestrian ways and airfield runways |
| | | 42.11.20 | Construction works for motorways, roads, streets and other vehicular or pedestrian ways and airfield runways |
| 06.4 * | Street lighting | 84.12.13 | Administrative housing and community amenity services |
| 06.5 | R & D housing and community amenities | 72.19.29 | Other research and experimental development services in engineering and technology, except biotechnology |
| | | 72.20.19 | Research and experimental development services in other social sciences |
| 06.6 | Housing and community amenities n.e.c. | 84.12.13 | Administrative housing and community amenity services |
| | | 42.99.12 | Sport and recreation constructions |
| | | 42.99.22 | Construction works for structures of outdoor stadia and sport grounds |
| 07.1 | Medical products, appliances and equipment | 84.12.12 | Administrative health care services |
| | | 84.30.11 | Compulsory social security services concerning sickness, maternity or temporary disablement benefits |
| | | 33.13.12 | Repair and maintenance services of irradiation, electromedical and electrotherapeutic equipment |
| | | 33.20.41 | Installation services of professional medical machinery and precision and optical instruments |
| 07.2 | Outpatient services | 86.21.10 | General medical practice services |
| | | 86.22.11 | Analysis and interpretation services of medical images |
| | | 86.22.19 | Other specialist medical practice services |
| | | 86.23.11 | Orthodontic services |
| | | 86.23.19 | Other dental practice services |
| 07.3 | Hospital services | 86.10.11 | Hospital surgical services |
| | | 86.10.12 | Hospital gynecological and obstetrical services |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|---|----------|---|
| | | 86.10.13 | Hospital rehabilitation services |
| | | 86.10.14 | Hospital psychiatric services |
| | | 86.10.15 | Other hospital services provided by medical doctors |
| | | 86.10.19 | Other hospital services |
| 07.4 | Public health services | 86.90.11 | Pregnancy-related services |
| | | 86.90.12 | Nursing services |
| | | 86.90.13 | Physiotherapeutic services |
| | | 86.90.14 | Ambulance services |
| | | 86.90.15 | Medical laboratory services |
| | | 86.90.16 | Blood, sperm and transplant organ bank services |
| | | 86.90.17 | Diagnostic imaging services without interpretation |
| | | 86.90.18 | Mental health services |
| | | 86.90.19 | Other human health services n.e.c. |
| | | 87.10.10 | Residential nursing care services |
| | | 87.20.11 | Residential care services for children suffering from mental retardation, mental health illnesses and substance abuse |
| | | 87.20.12 | Residential care services for adults suffering from mental retardation, mental health illnesses and substance abuse |
| | | 87.30.11 | Welfare services delivered through residential institutions to elderly persons |
| | | 87.30.12 | Welfare services delivered through residential institutions to disabled children and young people |
| | | 87.30.13 | Welfare services delivered through residential institutions to disabled adults |
| | | 87.90.11 | Other social work services with accommodation for children and young people |
| | | 87.90.12 | Social work services with accommodation for mistreated women |
| | | 87.90.13 | Other social work services with accommodation for adults |
| 07.5 | R & D Health | 72.19.30 | Research and experimental development services in medical sciences |
| 07.6 | Health n.e.c. | 84.12.12 | Administrative health care services |
| | | 75.00.12 | Veterinary services for livestock |
| 08.1 * | Administration of recreational and sporting affairs | 84.12.14 | Administrative recreational, cultural and religious services |
| 08.1 * | Operation of recreational and sporting facilities | 55.30.12 | Recreational and vacation camp services |
| | | 92.00.13 | Lotteries, numerical games and bingo services |
| | | 93.11.10 | Sports facility operation services |
| | | 93.13.10 | Services of fitness facilities |
| | | 93.19.11 | Sports and recreational sports event promotion services |
| | | 93.19.13 | Support services related to sports and recreation |
| | | 93.19.19 | Other sports and recreational sports services |
| | | 93.29.19 | Miscellaneous recreational services n.e.c. |
| | | 93.29.29 | Entertainment services n.e.c. |
| 08.2 * | Administration of cultural affairs | 84.12.14 | Administrative recreational, cultural and religious services |
| 08.2 * | Operation of cultural facilities | 90.02.11 | Performing arts event production and presentation services |
| | | 90.02.12 | Performing arts event promotion and organisation services |
| | | 90.02.19 | Other performing arts support services |
| | | 90.04.10 | Arts facility operation services |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|--|----------|---|
| | | 91.01.11 | Library services |
| | | 91.01.12 | Archive services |
| | | 91.02.10 | Museum operation services |
| | | 91.02.20 | Museum collections |
| | | 91.03.10 | Operation services of historical sites and buildings and similar visitor attractions |
| | | 91.04.11 | Botanical and zoological garden services |
| | | 91.04.12 | Nature reserves services, including wildlife preservation services |
| 08.3 * | Operation of publishing services | 18.12.11 | Printing services for postage stamps, taxation stamps, documents of titles, smart cards, cheques and other security papers and the like |
| | | 18.12.14 | Printing services for books, maps, hydrographic or similar charts of all kinds, pictures, designs and photographs, postcards |
| | | 58.11.11 | Printed educational textbooks |
| | | 58.11.12 | Printed professional, technical and scholarly books |
| | | 58.11.13 | Printed children books |
| | | 58.11.14 | Printed dictionaries and encyclopaedias |
| | | 58.11.15 | Printed atlases and other books with maps |
| | | 58.11.16 | Printed maps and hydrographic or similar charts, other than in book form |
| | | 58.11.19 | Other printed books, brochures, leaflets and the like |
| | | 58.11.20 | Books on disk, tape or other physical media |
| | | 58.11.30 | Online books |
| | | 58.12.10 | Directories and mailing lists printed or on physical media |
| | | 58.12.20 | Online directories and mailing lists |
| | | 58.12.30 | Licensing services for the right to use directories and mailing lists |
| | | 58.14.19 | Other printed journals and periodicals |
| | | 58.14.20 | Online journals and periodicals |
| | | 58.19.14 | Printed unused postage, revenue or similar stamps; stamp-impressed paper; cheque forms; banknotes, stock, share or bond certificates and similar documents of title |
| | | 58.19.19 | Other printed matter |
| | | 58.19.29 | Other online content n.e.c. |
| | | 58.29.29 | Other application software, packaged |
| | | 58.29.32 | Application software downloads |
| | | 58.29.40 | Online software |
| | | 58.29.50 | Licensing services for the right to use computer software |
| 08.3 * | Administration of broadcasting and publishing services | 84.13.14 | Administrative transport- and communications-related services |
| 08.3 * | Operation of broadcasting services | 60.10.11 | Radio programming and broadcasting services |
| | | 60.10.12 | Radio broadcasting originals |
| | | 60.10.20 | Radio channel programmes |
| | | 60.10.30 | Radio advertising time |
| | | 60.20.11 | Online television programming and broadcasting services, except by subscription |
| | | 60.20.12 | Other television programming and broadcasting services, except by subscription |
| | | 60.20.13 | Online television subscription programming and broadcasting services |
| | | 60.20.14 | Other television subscription programming and broadcasting services |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|--|----------|--|
| | | 60.20.20 | Television broadcasting originals |
| | | 60.20.31 | Television channel programmes, except for subscription television |
| | | 60.20.32 | Subscription television channel programmes |
| | | 60.20.40 | Television advertising time |
| 08.4 * | Administration of religious and other community services | 84.12.14 | Administrative recreational, cultural and religious services |
| 08.4 * | Operation of religious services | 94.91.10 | Services furnished by religious organisations |
| 08.4 * | Operation of other community services | 94.99.11 | Services furnished by human rights organisations |
| | | 94.99.12 | Services furnished by environmental advocacy groups |
| | | 94.99.13 | Special group protection services |
| | | 94.99.14 | Other civic betterment and community facility support services |
| | | 94.99.15 | Services provided by youth associations |
| | | 94.99.16 | Services provided by cultural and recreational associations |
| | | 94.99.17 | Services provided by other civic and social organisations |
| | | 94.99.19 | Services provided by other membership organisations n.e.c. |
| | | 94.99.20 | Grant-giving services by membership organisations |
| 08.5 | R & D recreation, culture and religion | 72.20.19 | Research and experimental development services in other social sciences |
| 08.6 | Recreation, culture and religion n.e.c. | 84.12.14 | Administrative recreational, cultural and religious services |
| 09.1 * | Inspection of schools providing pre-primary and primary education | 84.12.11 | Administrative educational services |
| 09.1 * | Provision of pre-primary and primary education | 85.10.10 | Pre-primary education services |
| | | 85.20.11 | Online primary education services |
| | | 85.20.12 | Other primary education services |
| 09.2 * | Inspection of schools providing secondary education | 84.12.11 | Administrative educational services |
| 09.2 * | Provision of secondary education | 85.31.11 | Online lower general secondary education services |
| | | 85.31.12 | Other lower general secondary education services |
| | | 85.31.13 | Online upper general secondary education services |
| | | 85.31.14 | Other upper general secondary education services |
| | | 85.32.11 | Online technical and vocational lower secondary education services |
| | | 85.32.12 | Other technical and vocational lower secondary education services |
| | | 85.32.13 | Online technical and vocational upper secondary education services |
| | | 85.32.14 | Other technical and vocational upper secondary education services |
| 09.3 * | Inspection of institutions providing post-secondary non-tertiary education | 84.12.11 | Administrative educational services |
| 09.3 * | Provision of post-secondary non-tertiary education | 85.41.11 | Online post-secondary non-tertiary general education services |
| | | 85.41.12 | Other post-secondary non-tertiary general education services |
| | | 85.41.13 | Online post-secondary non-tertiary technical and vocational education services |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|--|----------|--|
| | | 85.41.14 | Other post-secondary non-tertiary technical and vocational education services |
| 09.4 * | Inspection of universities and other institutions providing tertiary education | 84.12.11 | Administrative educational services |
| 09.4 * | Provision of tertiary education | 85.42.11 | Online first stage tertiary education services |
| | | 85.42.12 | Other first stage tertiary education services |
| | | 85.42.13 | Online second stage tertiary education services |
| | | 85.42.14 | Other second stage tertiary education services |
| | | 85.42.15 | Online third stage tertiary education services |
| | | 85.42.16 | Other third stage tertiary education services |
| 09.5 * | Inspection of institutions providing education not definable by level | 84.12.11 | Administrative educational services |
| 09.5 * | Provision of education not definable by level | 85.32.11 | Online technical and vocational lower secondary education services |
| | | 85.32.12 | Other technical and vocational lower secondary education services |
| | | 85.32.13 | Online technical and vocational upper secondary education services |
| | | 85.32.14 | Other technical and vocational upper secondary education services |
| | | 85.51.10 | Sports and recreation education services |
| | | 85.52.11 | Dancing schools and dance instructors services |
| | | 85.52.12 | Music schools and music instructors services |
| | | 85.52.13 | Fine arts schools and arts instruction services |
| | | 85.52.19 | Other cultural education services |
| | | 85.59.11 | Language school services |
| | | 85.59.12 | IT school services |
| | | 85.59.13 | Vocational education services n.e.c. |
| | | 85.59.19 | Education services n.e.c. |
| 09.6 * | Inspection of subsidiary services to education | 84.12.11 | Administrative educational services |
| 09.6 * | Subsidiary services to education | 85.60.10 | Educational support services |
| | | 55.90.11 | Room or unit accommodation services for students in student residences and school dormitories |
| 09.7 | R & D Education | 72.20.19 | Research and experimental development services in other social sciences |
| 09.8 | Education n.e.c. | 84.12.11 | Administrative educational services |
| 10.1 * | Sickness and disability (except benefits in kind) | 84.30.11 | Compulsory social security services concerning sickness, maternity or temporary disablement benefits |
| 10.1 * | Benefits in kind concerning sickness and disability | 86.1 | Hospital services |
| | | 86.2 | Medical and dental practice services |
| | | 86.9 | Other human health services |
| | | 88.10.13 | Vocational rehabilitation services for persons with disabilities |
| | | 88.10.14 | Visiting and assistance services for persons with disabilities |
| | | 88.10.15 | Day-care centre services for disabled adults |
| | | 88.91.11 | Child day-care services excluding day-care services for the disabled |
| | | 88.91.12 | Day-care services for disabled children and young people |

| COFOG | COFOG definition | CPA | CPA definition |
|--------|--|----------|--|
| 10.2 * | Old age (except benefits in kind) | 84.30.12 | Compulsory social security services concerning government employee pension schemes; old-age, disability or survivors' benefits other than for government employees |
| 10.2 * | Benefits in kind concerning old age | 88.10.11 | Visiting and assistance services for the elderly |
| | | 88.10.12 | Day-care centre services for the elderly |
| 10.3 * | Survivors (except benefits in kind) | 84.30.12 | Compulsory social security services concerning government employee pension schemes; old-age, disability or survivors' benefits other than for government employees |
| 10.3 * | Benefits in kind concerning survivors | 88.99.12 | Welfare services without accommodation |
| | | 88.99.19 | Other social services without accommodation n.e.c. |
| 10.4 * | Family and children (except benefits in kind) | 84.30.14 | Compulsory social security services concerning family and child allowances |
| 10.4 * | Benefits in kind concerning family and children | 88.91.13 | Baby-sitting services |
| | | 88.99.11 | Guidance and counselling services n.e.c. related to children |
| 10.5 * | Unemployment (except benefits in kind) | 84.30.13 | Compulsory social security services concerning unemployment compensation benefits |
| 10.5 * | Benefits in kind concerning unemployment | 78.10.12 | Permanent placement services, other than executive search services |
| | | 88.99.13 | Vocational rehabilitation services for the unemployed |
| 10.6 * | Provision of low-cost or social housing | 41.00.10 | Residential buildings |
| | | 68.20.11 | Rental and operating services of own or leased residential real estate |
| 10.6 * | Housing (except provision of low-cost or social housing) | 84.30.1 | Compulsory social security services |
| 10.7 * | Social exclusion n.e.c. (except benefits in kind) | 84.30.1 | Compulsory social security services |
| 10.7 * | Benefits in kind concerning social exclusion n.e.c. | 88.99.12 | Welfare services without accommodation |
| | | 88.99.19 | Other social services without accommodation n.e.c. |
| 10.8 | R & D Social protection | 72.20.19 | Research and experimental development services in other social sciences |
| 10.9 | Social protection n.e.c. | 84.11.13 | Overall economic and social planning and statistical services |

Annex 2: User manual

A user manual was prepared with the aim of enabling researchers and officers of a statistical institute or a research institute to replicate and expand the exercise developed as pilot test by the methodological study.

This document describes the installation process, how to configure the settings before running, how to run, and how to interpret and process the results.

Once the steps described in the process (Ch. 6) are completed — (i) identify and select the public service; (ii) build the ladder's rungs, i.e. the innovativeness levels of the service; (iii) define the tools: keyword list, baskets and basket combinations, URL list; the actions guided in the following pages — installation, setting, starting the program, results from the web crawling, processing the results — allow to start the real running of the web-scraping.

Installation

1: Install Node and Npm

With installing Node, you automatically install npm. The latest LTS version <https://nodejs.org/en/download> If the latest version is incompatible: the versions used in this project are:

[v8.1.4/](#)

[v8.11.3/](#)

2: Check Node and Npm installations

Check if both of these commands work. Go to the command line with WIN+R, type cmd, press enter, and type these commands:

- [Npm -v \[enter\]](#)
- [Node -v \[enter\]](#)

If one of these does not work, first solve the occurring issues before continuing. Use the knowledge at nodejs.org and google to solve.

3: Download program

At <https://github.com/wouterdialogic/urbino-data-acquirer> via the green Download link

4: Relocate download

Put the contents of 3 in an easy accessible directory (fe: c:\urbino)

5: Download additional files

Go to the command line, navigate to this folder, type npm install [enter]. This will download all the files needed.

Settings

These are the files you can edit:

1 **SETTINGS.JS**

- o Which file to use for keywords
- o Which file to use for URLs
- o Cutoff multiplier, its default value is 1.2. Change this only if you know how this works.

2 **[variable name] URLs.js**

You can change, add and remove URLs to be crawled in this file. Make sure to use this format.

[‘http://www.asrem.org’](http://www.asrem.org),

If you are starting a new URL, copy and paste this file and give it a proper name.

3 **[variable name] keywords.js**

You can change, add and remove keywords to be searched with in this file. Make sure to use this format.

‘codice identificativo’: 1,

The score (1 in this case) is not used at the moment, but added for later improvements to the program. If you are starting a new URL, copy and paste this file and give it a proper name.

Starting the program

In the command line prompt, go to the location where run.js is, and type: node run

Overall runtime depends on the number of URLs and the number of search-words.

After pressing enter the script starts and the following message is shown:

```
> node run
Starting with: http://portale.asl.at.it/
http://portale.asl.at.it/
Hi there, on this page we`ll provide information about our scraping progression.
```

Also, you will see continuous updates like this:

```
Starting with: http://www.asl.novara.it
http://www.asl.novara.it
Hi there, on this page we`ll provide information about our scraping progression.

We`ve scraped the mainpage: http://www.asl.novara.it
We`ll evaluate 22 pages:

1/22 http://www.asl.novara.it/indexrss_01.xml
2/22 http://www.asl.novara.it/wfprog/geturl.exe?id=3848&type=2
```

Due to the variety of links, it is not uncommon to see an error. This is not a critical issue: the program will continue to run even if it cannot process a certain webpage. For example, when the link refers to a word or pdf-file (which are not being parsed) the following screen is shown:

```
12/22 http://www.asl.novara.it/wfprog/geturl.exe?id=3389&type=2
Info: this is a link that cant be crawled!
http://www.asl.novara.it/wfprog/geturl.exe?id=3389&type=2
Out of respect we`ll skip this page and continue...
```

The program is done when the following message is shown.

```
No more municipalities to crawl!

Done!
```

Now you can see the results

Results from the web crawling

Tip

If you want to see the results while the program is running DO NOT open the csv file, since this will lock the file and the program will not be able to add records. Instead, make a copy of the csv file and open the copy. After inspecting, you can safely delete the copy for a clean workspace.

Per run 3 files are created:

```
[timestamp] scores urbino.csv
[timestamp] detailed results urbino.csv
[timestamp] overview results urbino.csv
```

All files contain a timestamp, for example: 6-19-2018_15-38-15 scores urbino.csv

```
[timestamp] scores urbino.csv
```

These are the 'settings' that are defined for a domain during run time. Let's say that for <http://www.amsterdam.nl> there are 30 pages visited via the main page. Those 30 visited pages

determine the score future crawled pages must have for this domain. A small example: If on 30 pages an average of 10 keywords are found, and the cut-off multiplier is set on 1.2. It will only evaluate pages that have $(10 * 1.2 =)$ at least 12 search-words found on them. In this file, the settings that determine further crawler are saved for evaluation and further tweaking.

`[timestamp] results urbino.csv`

This file contains all the URLs that have been visited that have a score that is higher than the `cutoff_multiplier` that is set for that domain. Each row is an URL, each column indicates how many and which keywords are found (an empty row is 0 keywords found)

`[timestamp] overview results urbino.csv`

This file contains a summary of the file detailed results Urbino.csv, all the scores are added for the domain. So each row represents a domain.

Processing results

When the web-scraping is complete a csv (comma separated values) file is created named `[timestamp] results_urbino.csv`. Copy the file (Ctrl C).

Open the Excel analys file. The Excel analysis file contains six tabs, namely: `Source`, `recoded`, `by subdomain`, `by domain`, `results`, and `Lookup keyword -> basket`.

Open the tab `Source`. Paste the results from the csv file into this tab. You can do this either by pasting it (Ctrl V) in the top left cel (A1) or by using the native Excel import wizard (File Open)

Open the tab `Lookup keyword -> basket`. In this tab the mapping between the keywords and the (sub)baskets is specified. Note: Make sure that the keywords are IDENTICALLY SPELLED to the keywords that have been used in the web-scraping (and that are now copied to the tab `Source`).

Open the tab `recoded`. Check the first row to see if the lookup is successful by looking at the column from D till the final column of the tab `Source+1`. The first row should now contain the baskets instead of the keywords. If not, manually adjust the keywords spelling list so that the lookup is successful. Note that sometimes there are spaces around the keywords. Excel does not automatically remove such spaces. This has been done manually.

Open the tab `by subdomain`. In this tab the data is aggregated to the basket level. Check if all the baskets are named in the first row. If not, add them manually. At this stage baskets were linked each other to determine the basket combinations.

Finally, open the tab `by domain`. In this tab the data is aggregated by basket and domain level. In this stage the service score is calculated too. By using a hierarchical `if-then` statement, the highest service score is assigned to each domain. In the tab `results` the distribution of the scores is shown; how many domains received the highest score, etc.

Help desk

Having experienced within the project team the replication of the procedures needed to prepare, run and complete the web-scraping process when extending it to additional sets of institutional websites, the setup of an off-line contact point is strongly recommended. The contact point should provide an effective service of answering questions and problem solving, accessible to those who will start experimenting with the methodology. As a minimum time-consuming activity is foreseen, it could be fulfilled by a subset of the StarPIN team during the experimental phase of application and refinement of the methodology. On a larger and systematic scale, it could become a service to be provided, as well as an area of exchange of best practice between researchers and statistical institutes.

Annex 3: Keywords and keyword baskets

Baseline keyword list for municipal waste collection pilot

| Level | Types of collection | Description | English boolean search(es) | Italian boolean search(es) | Dutch boolean search(es) |
|-------|-------------------------|---|---|--|--|
| 1 | No separate collection | Municipal solid waste are thrown into common bins in the street | Waste collection NOT recycling NOT separate collection | Raccolta rifiuti NOT ricicl* NOT raccolta differenziata | Afval OR vuilnis |
| 2 | Separate collection | Separate collection of waste (in common bins in the street) to recycle waste material | Waste collection AND separate collection AND recycling NOT door-to-door | Raccolta rifiuti AND raccolta differenziata AND ricicl* NOT porta-a-porta | Gescheiden afval AND ondergrondse afvalcontainer OR ondergrondse vuilcontainer |
| 3 | Door-to-door collection | Separate waste streams are collected in separate bins directly at home (periodically) | Waste collection AND separate collection AND recycling AND door-to-door | Raccolta rifiuti AND raccolta differenziata AND ricicl* AND porta-a-porta | Kliko OR groencontainer |
| 4 | Pay-as-you-throw | Separate waste streams are collected in separate bins directly at home (periodically), while a tariff (proportional to the weight) is applied to unsorted waste | Waste collection AND separate collection AND recycling AND (tariff OR pay-as-you-throw) | Raccolta rifiuti AND raccolta differenziata AND ricicl* AND tariffa puntuale | Afvalpas OR huisvuilpas OR {toegangspas NEAR container} |

Final keyword list and baskets for the municipal waste collection pilot service: English, Italian, Dutch

| Basket | Keyword_ENGLISH | Keyword_ITALIAN | Keyword_DUTCH |
|--------|------------------------------------|--|---|
| 1a | daily collection | raccolta giornaliera | dagelijkse inzameling |
| 1a | fixed tariff AND waste | tariffa rifiuti | vast tarief AND afval |
| 1a | garbage | spazzatura | vuilnis |
| 1a | garbage collection | raccolta dei rifiuti | afvalinzameling |
| 1a | household garbage | rifiuti domestici | huisvuil |
| 1a | refuse manager | gestore rifiuti | afvalbeheerder |
| 1a | trash bin | cestino dei rifiuti | afvalbak |
| 1a | waste | rifiut* | afval |
| 1a | waste collection | raccolta rifiuti | afvalpunt |
| 1b | number of residents | numero residenti | aantal bewoners |
| 1b | size household | dimensioni | grootte huishouden |
| 1b | size household | metratura | grootte huishouden |
| 1b | size household | superficie | grootte huishouden |
| 1c | bulky waste | rifiuti ingombranti | grofvuil |
| 1c | debris | detriti | puin |
| 3a | at home | a domicilio | aan huis |
| 3a | at home | domiciliare | aan huis |
| 3a | collection days | giorni di raccolta | ophaaldagen |
| 3a | colored bins | cassonetti colorati | gescheiden afvalcontainer |
| 3a | Common bins | bidoni comuni | straatcontainer |
| 3a | common bins | cassonetti comuni | straatcontainer |
| 3a | condominium bins | bidoni condominiali | straatcontainer |
| 3a | condominium bins | cassonetti condominiali | straatcontainer |
| 3a | door to door | porta a porta | huis aan huis |
| 3a | door-to-door | porta-a-porta | huis-aan-huis |
| 3a | waste calendar | calendario dei rifiuti | afvalkalender |
| 3a | waste schedule | programma dei rifiuti | ophaaldagen |
| 3b | biodegradable bag | sacchetto biodegradabile | biologisch afbreekbare zak |
| 3b | biodegradable bag | sacchetto biodegradabile | GFT-zak |
| 3b | drinks carton | cartone | PMD |
| 3b | ecological wet bag | sacchetto ecologico umido | biologisch afbreekbare zak |
| 3b | ecological wet bag | sacchetto ecologico umido | biologisch afbreekbare zak |
| 3b | fruit, vegetables and garden waste | umido | groente, fruit en tuinafval |
| 3b | FVG | organico | GFT |
| 3b | glass bell | campana per il vetro | glasbak |
| 3b | glass container | campana vetro | glasbak |
| 3b | glass container | campana per il vetro | glascontainer |
| 3b | paper * | carta | papiercontainer |
| 3b | plastic * | plastica | plastic |
| 3b | plastic tin | plastica e lattine | plastic verpakkingen, blik en drinkpakken |
| 3b | recycling | ricicl* | recycling |
| 3b | recycling area | area di riciclo | sorteerinstallatie |
| 3b | residual waste | indifferenziato | restafval |
| 3b | residual waste | residuo, indifferenziato | restafval |
| 3b | separate collection | raccolta differenziata | afvalscheiding |
| 3b | wheelie bin | cassonetto | kliko |
| 4a | collection containers | contenitori per raccolta | verzamelcontainers |
| 4a | ecological station | stazione ecologica | ecologisch station |
| 4a | environmental station | stazione ecologica | milieubringstation |
| 4a | mobile waste separation station | stazione mobile di separazione rifiuti | mobiel scheidingsstation |

| Basket | Keyword_ENGLISH | Keyword_ITALIAN | Keyword_DUTCH |
|--------|------------------------------------|--|-----------------------------------|
| 4a | oat | sistema sotterraneo di smistamento rifiuti | oat |
| 4a | underground waste transport system | sistema di trasporto rifiuti sotterraneo | ondergronds afvaltransportsysteem |
| 4a | underground waste transport system | sistema sotterraneo di trasporto rifiuti | ondergronds-afvaltransportsysteem |
| 4a | underground container | contenitore sotterraneo | ondergrondse container |
| 4a | underground collection containers | contenitori per raccolta sotterranea | ondergrondse verzamelcontainers |
| 4a | underground waste container | contenitore per rifiuti sotterraneo | ondergrondse afvalcontainer |
| 4a | underground-collection containers | contenitore per raccolta sotterraneo | ondergrondse verzamelcontainers |
| 4a | waste collection point | punto di raccolta rifiuti | afvalbrengpunt |
| 4a | waste collection station | stazione di raccolta rifiuti | afvalbrengstation |
| 4a | waste collection station | stazione di raccolta rifiuti | kliko verzamelplaats |
| 4a | waste dump | discarica rifiuti | afvalbrenghaak |
| 4a | waste separation station | stazione di separazione rifiuti | afvalscheidingsstation |
| 4a | waste transport system | sistema di trasporto dei rifiuti | afvaltransportsysteem |
| 5a | bin | bidone | vuilnisbak |
| 5a | chip in container | chip nel contenitore | chip in container |
| 5a | chip in waste container | contenitore con microchip | chip in afvalcontainer |
| 5a | chip in waste container | contenitore dotato di microchip | chip van afvalcontainer |
| 5a | chip in wheelie bin | bidone con chip | chip in kliko |
| 5a | chip on container | chip sul contenitore | chip van container |
| 5a | chip* container | chip * contenitore | chippen van container |
| 5a | chip* container | contenitore * chip | chippen containers |
| 5a | container chip | contenitore con chip | container chip |
| 5a | containerchip | bidone con chip | containerchip |
| 5a | mini container | contenitore | minicontainer |
| 5a | personal container at home | contenitore personale a domicilio | kliko |
| 5a | waste container chip | contenitore per rifiuti con chip | afvalcontainer chip |
| 5a | waste-container-chip | contenitore per rifiuti con chip | afvalcontainerchip |
| 5a | wheelie bin chip | bidone con chip | kliko chip |
| 5a | wheelie bin chip | bidone con chip | chip van kliko |
| 5a | wheelie-bin-chip | bidone con chip | klikochip |
| 5b | bag counting | conteggio sacchetti | |
| 5b | bag number | numero sacchetti | |
| 5b | number of bags | numero dei sacchetti | aantal vuilniszakken |
| 5b | volume | volume | volume |
| 5c | chip | chip | chip |
| 5c | microchip | microchip | microchip |
| 6a | access card | tessera personalizzata | toegangspas |
| 6a | environmental pass | tessera ambiente | milieupas |
| 6a | identification code | codice identificativo | identificatiecode |
| 6a | identification microchip | microchip indentificativo | identificatie microchip |
| 6a | identification microchip | microchip indentificativo | identificatie microchip |
| 6a | key | chiavetta | sleutel |
| 6a | magstripe OR swipe card | tessera magnetica | magneetstrip |
| 6a | personalized card | tessera personalizzata | gepersonaliseerde kaart |

| Basket | Keyword_ENGLISH | Keyword_ITALIAN | Keyword_DUTCH |
|--------|-----------------------------|-------------------------------|---------------------------------|
| 6a | personalized key | chiavetta personalizzata | gepersonaliseerde sleutel |
| 6a | waste pass | tessera personalizzata | afvalpas |
| 6b | pay-as-you-throw | tariffa puntuale | diftar |
| 6b | pay-as-you-throw | tariffa puntuale | pay-as-you-throw |
| 6b | variable rate of the tariff | quota variabile della tariffa | variabele tarief van het tarief |
| 6b | variable rate of the tariff | quota variabile della tariffa | variabele tarief van het tarief |
| 6b | variable tariff | tariffa variabile | variabel tarief |
| 6c | automating weighing | pesatura automatica | automatische weging |
| 6c | bag weighing | pesa dei sacchetti | afvalzak wegen |
| 6c | weighing | pesatura | weging |
| 6c | weight | peso | gewicht |

Final keyword list and baskets for the municipal waste collection pilot service: Portuguese, Slovak, Swedish

| Basket | Keyword_PORTUGUESE | Keyword_SLOVAKIAN | Keyword_SWEDISH |
|--------|--|--|--|
| 1a | lixreira | odpadkový kôš | sopkorg, soptunna |
| 1a | gerente de recusas | manažér v odpadovom hospodárstve | sopansvarig, ansvarig för sopor |
| 1a | recolha de lixo | zber odpadkov | sopinsamling, avfallsinsamling |
| 1a | recolha diária | denný zber | daglig sopinsamling, daglig avfallsinsamling |
| 1a | recolha de resíduos/lixos | zber odpadu | sopinsamling, avfallsinsamling |
| 1a | lixo | odpad | avfall, sopor |
| 1a | tarifa lixo/taxa de esgotos | pevná tarifa A odpad | fast taxa och avfall |
| 1a | lixo | odpadky | sopor, avfall |
| 1a | lixo doméstico | odpadky z domácností | hushållsavfall |
| 1b | dimensão da família | veľkosť domácnosti | hushållsstorlek, storlek av hushåll |
| 1b | dimensão da família | veľkosť domácnosti | hushållsstorlek, storlek av hushåll |
| 1b | número de moradores | počet obyvateľov | antal boende |
| 1b | dimensão da família | veľkosť domácnosti | hushållsstorlek, storlek av hushåll |
| 1c | detritos | trosky/úlomky | massor, fyllmassor, avfall, debris |
| 1c | resíduos volumosos | objemný odpad | skrymmande avfall |
| 3a | em casa | doma | hemma |
| 3a | caixotes comuns | spoločné zberné nádoby na odpad | allmänna behållare, allmänna soptunnor |
| 3a | caixotes de condomínio | kondomíniové koše | föreningens/husets avfallsbehållare/soptunnor/avfalls kasuner/sopkasuner |
| 3a | calendário de resíduos | kalendár zberu odpadu | sophämtningsskalendern, körschema |
| 3a | caixotes coloridos/ caixotes de reciclagem | farebné zberné nádoby na odpad | färgade soptunnor |
| 3a | caixotes comuns | spoločné zberné nádoby na odpad | allmänna behållare, allmänna soptunnor |
| 3a | caixotes de condomínio | kondomíniové koše | föreningens/husets avfallsbehållare/soptunnor/avfalls kasuner/sopkasuner |
| 3a | em casa | doma | hemma |
| 3a | dias de recolha | dni zberu opadu | insamlingsdagar |
| 3a | de porta em porta | od dverí k dverám | dörr till dörr |
| 3a | de porta em porta | od dverí k dverám | dörr till dörr |
| 3a | agenda de resíduos/ calendário de recolha | harmonogram zberu odpadov | avfallsschemat, sophämtningsschema |
| 3b | saco molhado ecológico | ekologické vodeodolné vrecká | ekologisk matavfallspåse |
| 3b | saco biodegradável | biologicky rozložiteľné vrecká/sáčky | biologiskt nedbrytbar påse |
| 3b | saco biodegradável | biologicky rozložiteľné vrecká/sáčky | biologiskt nedbrytbar påse |
| 3b | saco molhado ecológico | ekologické vodeodolné vrecká | ekologisk matavfallspåse |
| 3b | área de reciclagem | miesto vyhradené pre recyklovanie odpadu | återvinningsområde |
| 3b | recipiente de vidro | sklenená nádoba | glasbehållare |
| 3b | papel * | papier * | papper * |
| 3b | embalagem de papelão de bebidas | nápojový kartón | pappersförpackningar, wellpapp |
| 3b | latão | nádoby na odpad s kolieskami | sopkärl, avfallskasun |

| Basket | Keyword_PORTUGUESE | Keyword_SLOVAKIAN | Keyword_SWEDISH |
|--------|--|--------------------------------------|--|
| 3b | vidrão | sklenený zvon | glasbehållare |
| 3b | lixos residuais | zvyškový odpad | restavfall |
| 3b | lixo orgânico | odpad z ovocia, zeleniny a záhrad | matavfall |
| 3b | plástico * | plast * | plast * |
| 3b | recolha separada | triedený zber odpadu | separat insamling |
| 3b | lixos residuais | zvyškový odpad | restavfall |
| 3b | reciclagem | recyklácia | återvinning |
| 3b | saco biodegradável | biologicky rozložiteľné vrecká/sáčky | biologiskt nedbrytbar påse |
| 3b | restos de frutas, legumes e verduras | odpad z ovocia, zeleniny a záhrad | Frukt- grön- och trädgårdsavfall |
| 3b | recipiente por vidro | sklenená nádoba | glasbehållare |
| 3b | lata e plástico | plastová nádoba | behållare av hårdplast och burkar |
| 4a | recipientes de recolha subterrânea | kontajner na podzemný zber odpadu | underjordiska uppsamlingsbehållare/kasuner |
| 4a | contentor de resíduos subterrâneo | podzemný kontajner na odpad | underjordisk avfallsbehållare/kasuner |
| 4a | contentor subterrâneo | podzemný kontajner | underjordisk behållare/kasuner |
| 4a | recipientes de recolha subterrânea | kontajner na podzemný zber odpadu | underjordiska uppsamlingsbehållare/kasuner |
| 4a | recipientes de recolha | zberné nádoby | uppsamlingsbehållare/kasuner/kär l/tunnor |
| 4a | despejo de lixo | skládka odpadu | avfallsställe/avfallsområde |
| 4a | ponto de recolha de resíduos | zberné miesto | avfalls insamlingsplats |
| 4a | sistema de transporte de resíduos | systém prepravy odpadu | avfalls transportsystem |
| 4a | sistema subterrâneo de transporte de resíduos | podzemný systém prepravy odpadu | underjords avfallstransportsystem |
| 4a | | ovos | underjords avfallshanteringssystem |
| 4a | sistema de transporte de resíduos subterrâneos | podzemný systém prepravy odpadu | underjords-avfalls-transport-systemet |
| 4a | estação de recolha de resíduos | prekládková stanica pre odpad | återvinningscentral, återvinningsstation, avfallsinsamlingsstation |
| 4a | estação de separação de resíduos | miesto na triedenie odpadu | avfalls sorteringsstation/sorteringscentral |
| 4a | estação ambiental | environmentálna stanica | miljö station |
| 4a | estação móvel de separação de resíduos | mobilný zber triedeného odpadu | mobil miljöstation |
| 4a | estação ecológica | ekologická stanica | ekologisk miljöstation |
| 5a | mini contentor | mini kontajner | mini container, minicontainer |
| 5a | contentor com chip | containerchip | container med chip |
| 5a | contentor com chip | čip * kontajner | chip * behållare/kärl |
| 5a | chip no contentor | čip v kontajneri | chip i behållaren/kärlet |
| 5a | chip no contentor | čip na kontajneri | chip på behållaren/kärlet |
| 5a | contentor com chip | kontajnerový čip | behållare med chip |
| 5a | chip em contentor do lixo | čip v kontajneri na odpad | chip i avfallsbehållare/kärl |
| 5a | contentor do lixo com chip | waste-container-chip | avfallscontainer/kärl med chip |
| 5a | contentor pessoal em casa | vlastný kontajner doma | personlig avfallsbehållare/kärl hemma |
| 5a | contentor de latão com chip | wheelie-bin-chip | sopkärl-chip |

| Basket | Keyword_PORTUGUESE | Keyword_SLOVAKIAN | Keyword_SWEDISH |
|--------|---|--------------------------------|---|
| 5a | contentor de latão com chip | čip pre kolieskový kontajner | sopkärlschip |
| 5a | chip em contentor de latão | čip v kolieskovom kontajneri | chip i sopkärl |
| 5a | chip de contentor de latão | čip do kolieskového kontajneru | sopkärlschip |
| 5a | chip * contentor | čip * kontajner | chip * avfallsbehållare/container/kärl |
| 5a | chip em contentor de resíduos | čip v kontajneri na odpad | chip i avfallsbehållare/kärl |
| 5a | chip de contentor de resíduos | čip do kontajneru na odpad | avfallsbehållare med chip |
| 5a | caixote | nádoba na odpad/odpadkový kôš | sopkärl |
| 5b | contagem de saco | počítanie vreciek | påsräkning |
| 5b | volume | objem | volym |
| 5b | número do saco | číslo vrecka | påsnummer |
| 5b | número de sacos | počet vreciek | antal påsar |
| 6a | chave personalizada | personalizovaný kľúč | personlig nyckel |
| 6a | chave | kľúč | nyckel |
| 6a | Código de identificação | identifikačný kód | identifieringskod |
| 6a | microchip de identificação | identifikačný mikročip | identifierande mikrochip |
| 5c | chip | čip | chip/chipp |
| 6a | cartão personalizado | personalizovaná karta | personligt kort |
| 6a | passé para resíduos | odpadový preukaz | avfallskort |
| 6a | passé ambiental | environmentálny preukaz | miljökort |
| 6a | cartão de acesso | prístupová karta | passerkort |
| 6a | microchip de identificação | identifikačný mikročip | identifierande mikrochip |
| 6a | barra magnética | magstripe OR swipe card | kort med magnetremsa ELLER passerkort |
| 6b | parte variável da tarifa | variabilná sadzba tarify | tariffens rörliga avgift |
| 6b | pay-as-you-throw/ pagamento contra entrega | platiť za to, čo vyhodíš | rörlig avgift |
| 6b | tarifa variável | variabilná tarifa | rörlig avgift |
| 6b | parte variável da tarifa | variabilná sadzba tarify | tariffens rörliga avgift |
| 6b | pay-as-you-throw/ pagamento contra entrega | platiť za to, čo vyhodíš | rörlig avgift |
| 6c | peso | váha/závažie | vikt |
| 6c | pesagem | váženie | vägning |
| 6c | pesagem automática | automatizované váženie | automatisering av vägning |
| 6c | peso do saco | váženie vreciek | påsvägning, vägning av påsar |
| 5c | microchip | mikročip | mikrochip |
| 5a | contentor (ponto de recolha) de electrodomésticos | | |
| 5a | contentor (ponto de recolha) de pilhas | | |
| 5a | contentor (ponto de recolha) de óleos usados | | |
| 5a | contentor (ponto de recolha) de plástico | | |

Keyword list for the Health homecare pilot service

| Types of homecare services | Description | English boolean search |
|-------------------------------|--|---|
| Traditional homecare services | Healthcare professionals (nurses) go to patients' homes and take care of them | Homecare AND social service AND nurse* AND/OR physiotherapist* AND/OR paramedical |
| Remote telemedicine services | Diagnosis and monitoring is undertaken by healthcare professionals at distance | Homecare AND telemedicine AND/OR remote control AND/OR remote monitoring NOT nurse* |
| e-Health services | Advanced systems of monitoring patients, interacting with them and pushing them to self-responsibility | e-health AND digital AND/OR personal assistant AND/OR auto-monitoring AND/OR empowerment NOT nurse* |

Final keyword list for the Health homecare pilot service

| Basket | Keywords_ENGLISH | Keywords_ITALIAN | Keywords_DUTCH |
|--------|-----------------------|----------------------------------|-----------------------------|
| 1a | care | cure | zorg |
| 1a | care | cura | zorg |
| 1a | care | assistenza | hulp |
| 1a | health care | assistenza sanitaria | gezondheidszorg |
| 1a | therapeutic education | educazione terapeutica | therapeutische activiteiten |
| 1b | domestic | domiciliare | thuis |
| 1b | domiciliary | domiciliarità | woonachtig |
| 1b | home care, home help | cure domiciliari | thuiszorg |
| 1b | homecare | assistenza domiciliare | thuiszorg |
| 1b | monitoring | monitoraggio | monitoring |
| 1b | residence | domicilio | woning |
| 2b | auto monitoring | auto monitoraggio | automatische monitoring |
| 2b | automonitoring | automonitoraggio | continue glucose meten |
| 2b | distance | distanza | op afstand |
| 2b | domotics | domotica | domotica |
| 2b | domotics | domotica | zorgdomotica |
| 2b | integrated home care | assistenza domiciliare integrata | integrale thuiszorg |
| 2b | integrated home care | ADI | integrale thuiszorg |
| 2b | remote control | controllo remoto | afstandbediening |
| 2b | sensoristic | sensoristica | |
| 2b | tele aid | tele soccorso | hulp op afstand |
| 2b | tele assistance | tele assistenza | zorg op afstand |
| 2b | tele consultation | tele consulto | teleconsultatie |
| 2b | tele control | tele controllo | zorg op afstand |
| 2b | tele health | tele salute | telehealth |
| 2b | tele medicine | tele medicina | telemedicine |
| 2b | tele monitoring | tele monitoraggio | telemonitoring |
| 2b | tele nursing | tele infermieristica | zorg op afstand |
| 2b | tele rehabilitation | tele riabilitazione | revalidatie op afstand |
| 2b | teleaid | telesoccorso | beeldzorg |
| 2b | tele-aid | tele-soccorso | hulp op afstand |
| 2b | teleassistance | teleassistenza | zorg op afstand |
| 2b | tele-assistance | tele-assistenza | zorgdomotica |
| 2b | teleconsultation | teleconsulto | teleconsultatie |
| 2b | tele-consultation | tele-consulto | teleconsult |

| Basket | Keywords_ENGLISH | Keywords_ITALIAN | Keywords_DUTCH |
|--------|----------------------------------|------------------------------|-------------------------------|
| 2b | telecontrol | telecontrollo | telezorg |
| 2b | tele-control | tele-controllo | telezorg |
| 2b | telehealth | telesalute | telehealth |
| 2b | tele-health | tele-salute | telecare |
| 2b | telehomecare | telehomecare | tele-homecare |
| 2b | telemedicine | telemedicina | telemedicine |
| 2b | tele-medicine | tele-medicina | telemedicine |
| 2b | telemonitoring | telemonitoraggio | thuis meten |
| 2b | tele-monitoring | tele-monitoraggio | monitoren op afstand |
| 2b | tele-nursing | tele-infermieristica | zorg op afstand |
| 2b | telerehabilitation | teleriabilitazione | revalidatie op afstand |
| 2b | tele-rehabilitation | tele-riabilitazione | revalidatie op afstand |
| 2b | teleconsult | televisita | teleconsult |
| 2b | | tediagnosi | diagnose op afstand |
| 2b | | tediagnostica | diagnose op afstand |
| 2c | auto-monitoring | auto-monitoraggio | continue glucosemonitoring |
| 2c | digital ECG | ECG digitale | digitale ECG |
| 2c | digital glucometer | glucometro digitale | digitale glucosemeter |
| 2c | digital glucometer | glucometro digitale | digitale bloedglucosemeter |
| 2c | digital stethoscope | stetoscopio digitale | digitale stethoscoop |
| 2c | remote monitoring blood pressure | | thuismeten bloeddruk |
| 2c | remote monitoring COPD | | thuismeten COPD |
| 2c | remote monitoring heart failure | | thuismeten hartfalen |
| 2c | tele cardiology | tele cardiologia | hartbewaking op afstand |
| 2c | tele dermatology | tele dermatologia | teledermatologie |
| 2c | tele diabetology | tele diabetologia | |
| 2c | tele radiology | tele radiologia | teleradiologie |
| 2c | telecardiology | telecardiologia | hartbewaking op afstand |
| 2c | tele-cardiology | tele-cardiologia | hartbewaking op afstand |
| 2c | teledermatology | teledermatologia | teledermatologie |
| 2c | tele-dermatology | tele-dermatologia | teledermatologie |
| 2c | tediabetology | tediabetologia | |
| 2c | tele-diabetology | tele-diabetologia | |
| 2c | telenursing | teleinfermieristica | zorg op afstand |
| 2c | teleradiology | teleradiologia | teleradiologie |
| 2c | tele-radiology | tele-radiologia | teleradiologie |
| 3a | interaction | interazione | interactie |
| 3a | interactive | interattivo | interactief |
| 3b | app | app | app |
| 3b | health technology assessment | health technology assessment | health technology assessment |
| 3b | HTA | HTA | HTA |
| 3b | iPad | iPad | iPad |
| 3b | mhealth | mhealth | mhealth |
| 3b | mobile | mobile | mobiel |
| 3b | smart health | smart health | smart health |
| 3b | smartphone | smartphone | smartphone |
| 3b | tablet | tablet | tablet |

Final keyword list for the Health homecare pilot service

| Basket | Keywords_PORTUGUESE | Keywords_SLOVACK | Keywords_SWEDEN |
|--------|--|-------------------------------------|---|
| 1a | casa | Domov | Hem |
| 1a | doméstica | domáci | inhemsk |
| 1a | residência | bydliska | bostad |
| 1a | domiciliar | domáci | hem-(-) |
| 1b | Cuidado | starostlivost' | vård |
| 1b | Cuidado | starostlivost' | vård |
| 1b | atendimento domiciliar, ajuda domiciliar | domáca starostlivost', domáca pomoc | hemvård, hemhjälp |
| 1b | enfermeira | zdravotná sestra | sjuusköterska |
| 1b | enfermeira | zdravotná sestra | sjuusköterska |
| 1b | enfermeiros | sestry | sjuusköterskor |
| 1b | paramédico | zdravotník | akutvårdare, ambulanspersonal, sjukvårdare |
| 1b | paramédico | zdravotník | akutvårdare, ambulanspersonal, sjukvårdare |
| 1b | paramédico | zdravotník | akutvårdare, ambulanspersonal, sjukvårdare |
| 1b | monitorizar | monitorovanie | övervakning |
| 1a | assistência domiciliária | domáca starostlivost' | vård i hemmet |
| 1a | residência | bydliska | bostad |
| 1b | cuidado | starostlivost' | vård |
| 1b | cuidados de saúde | zdravotná starostlivost' | sjukvård |
| 1b | fisioterapia | fyzioterapia | fysioterapi |
| 1b | fisioterapeuta | fyzioterapeut | fysioterapeut |
| 1b | educação terapêutica | terapeutické vzdelávanie | terapeut utbildning |
| 2c | telefone | telefónu | telefonenhet, telefonapparat |
| 2c | controle/controlo remoto | dial'kové ovládanie | fjärrkontroll |
| 2c | tele * | teľa * | tele * |
| 2a | tele monitorização (tele monitorizar) | tele monitorovanie | teleövervakning |
| 2a | tele-monitorização (tele-monitorizar) | tele-monitoring | tele-övervakning |
| 2a | telemonitorização (telemonitorizar) | telemonitorovanie | fjärrövervakning |
| 2a | tele saúde | tele zdravie | tele hälsa, e-hälsa |
| 2a | tele-saúde | tele-health | tele-hälsa, e-hälsa |
| 2a | telessaúde | Telehealth | telehälsa e-hälsa |
| 2a | tele medicina | tele medicína | tele medicin |
| 2a | tele-medicina | tele-medicína | tele-medicin |
| 2a | telemedicina | telemedicína | telemedicin |
| 2c | controle/controlo remoto | | fjärrkontroll |
| 2c | remoto | dial'kový | avlägsen, fjärr |
| 2a | monitorização automática | auto monitorovanie | automatisk övervakning, automatiserad övervakning |
| 2a | monitorização-automática | auto-monitoring | automatisk-övervakning |
| 2a | automonitorização | automonitoring | automatisk övervakning |

| Basket | Keywords_PORTUGUESE | Keywords_SLOVACK | Keywords_SWEDEN |
|--------|-------------------------------------|----------------------------------|--|
| 2a | telecontrolo/telecontrolo | diaľkové riadenie | Fj rrstyrning, telefonkontroll |
| 2a | tele controle/controlo | telekontrol | Fj rrstyrning, |
| 2a | tele-controle/controlo | tele-control | fj rrstyrning |
| 2a | teleconsulta | telekonzult cie | Distanskonsultation, telefonkonsultation |
| 2a | consulta tele/telef nica | tele konzult cie | konsultation p  distans |
| 2a | tele-consulta | tele-konzult cie | distans-konsultation |
| 2a | tele-reportagem/ telereportagem | tele-reporting | distans-rapportering, telefonrapport |
| 2a | tele-visita | tele-n v števa | distans-l karbes k, l karbes k p  distans, telefonbes k |
| 2a | tele-diagn stico | tele-diagnostika | diagnos p  distans, diagnos via telefon |
| 2a | tele-diagn stico | tele-diagnostika | diagnostik p  distans |
| 2c | dist ncia | vzdialenosť | distans |
| 2a | teleassist ncia | teleassistance | mobilsupport, fj rrassistans, fj rrhj lp, hj lp p  distans, telefonassistans |
| 2a | assist ncia tele | tele asistencie | mobilsupport, fj rrassistans, fj rrhj lp, hj lp p  distans |
| 2a | tele-assist ncia | tele-asistencia | mobilsupport, fj rrassistans, fj rrhj lp, hj lp p  distans |
| 2a | teleajuda/telesocorro | TeleAid | telehj lp, telest d, hj lv  ver telefon |
| 2a | tele-ajuda/tele socorro | tele pomoci | tele hj lp, tele st d |
| 2a | tele-ajuda/tele-socorro | tele-aid | tele-hj lp, tele-st d |
| 2c | dispositivo | zariadenie | anordning, apparat |
| 2c | instala  o | in stal cia | installation |
| 2a | cuidados domicili rios integrados | integrovan  dom ca starostlivosť | integrerad hemsjukv rd, integrerad hemv rd |
| 2a | Assist ncia domicili ria telef nica | telehomecare | hemv rd p  distans |
| 2a | teleassistido | teleassisted | Hemv rdad, telefonassistans |
| 2c | dispositivo | zariadenie | anordning, apparat |
| 2b | telecardiologia | telecardiology | kardiologi p  distans, telekardiologi |
| 2b | telerradiologia | teler diol gia | radiologi p  distans, teleradiologi |
| 2b | teledermatologia | teledermatol giu | dermatologi p  distans, teledermatologi |
| 2b | telerreabilita  o | telerehabilitation | rehabilitering p  distans |
| 2b | tele enfermagem | telenursing | sk tsel p  distans, omv rd p  distans |
| 2b | telediabetologia | telediabetology | diabetologi p  distans |
| 2b | tele cardiologia | tele kardiol gia | kardiologi p  distans |
| 2b | tele radiologia | r diol gia | radiologi p  distans |
| 2b | tele dermatologia | tele dermatol gia | dermatologi p  distans |
| 2b | tele reabilita  o | telehabilit cia | rehabilitering p  distans |

| Basket | Keywords_PORTUGUESE | Keywords_SLOVACK | Keywords_SWEDEN |
|--------|----------------------------------|---------------------------------------|--|
| 2b | tele enfermagem | tele ošetrovateľstvo | skötsel på distans, omvård på distans |
| 2b | tele diabetologia | dia diabetológia | diabetologi på distans |
| 2b | tele-cardiologia | tele-kardiológia | kardiologi på distans |
| 2b | tele-radiologia | tele-rádiológmi | radiologi på distans |
| 2b | tele-dermatologia | tele-dermatológia | dermatologi på distans |
| 2b | tele-reabilitação | tele-rehabilitačné | rehabilitering på distans |
| 2b | tele-enfermagem | tele-ošetrovateľskej | skötsel på distans, omvård på distans |
| 2b | tele-diabetologia | tele-diabetológia | diabetologi på distans, tele-diabetologi |
| 3c | interativo | interaktívne | interaktiv |
| 3c | interação | interakcia | interaktion |
| 3a | e-health/e-saúde | e-health | e-hälsa |
| 3a | ehealth/esaúde | eHealth | eHälsa |
| 3a | avaliação de tecnologia em saúde | hodnotenie zdravotníckych technológií | utvärdering av hälsoteknik |
| 3a | | | |
| 3c | aplicativo/app | aplikácie | app |
| 3b | programas | softvér | programvara |
| 3a | saúde inteligente | inteligentné zdravie | smart hälsa |
| 3c | digital | digitálne | digital |
| 3c | digital | digitálne | digital |
| 3c | fortalecimento | splnomocnení | Bemyndigande, möjliggörande |
| 3c | smart box | smart box | smartbox |
| 3c | smartphone | smartphone | smartphone |
| 3c | domótica | domotics | domotik |
| 3c | | sensoristic | sensorisk |
| 3b | eletrocardiograma digital | digitálny elektrokardiogram | digitalt elektrokardiogram |
| 3b | estetoscópio digital | digitálny stetoskop | digitalt stetoskop |
| 3b | glicosímetro digital | digitálny glukometr | digital glukometer |

Keyword baskets — For waste collection pilot service

The keywords selected for the waste collection pilot service were grouped in the following baskets:

1. waste, waste bins, waste collection (baseline characteristic: identifies that a waste management service exists);
2. 'structural' variables which influence the tariff (house size, no of inhabitants);
3. debris or bulky waste (not strictly related with the service evolution, but relevant in some countries, namely NL);
4. home collection based on waste 'fetching';
5. separate collection;
6. waste collection based on waste 'bringing';
7. home collection based on waste 'fetching' with personalised systems;
8. measurement of waste amount based on the number of bags/volume;

9. identification system based on information encoded in a chip;
10. personalised systems of waste 'bringing';
11. measurement of waste amount based on weight/weighing systems;
12. pay-as-you-throw, including tariff systems (variable tariff/tariff rate).

To capture the 'AND' operator, the coexistence of baskets which are found on the webpage is checked for in each of the pre-selected URLs. A basic combination was used to detect the level of innovativeness of a service and to attribute a level indicator to the institution providing that service. In our first pilot case on waste collection, the first set of combinations tentatively assign the innovativeness level as:

1 (2 adds details): the URL contains information related to the description of waste collection, with no further details on service characteristics = **innovation level 1**

4 and 5; evidence of separate waste collection and of garbage 'fetch' at home = **innovation level 2.a**

5 and 6: evidence of separate waste collection and of garbage 'bring' practices = **innovation level 2.b**

1 and 7, 1 and 8, 1 and 9: the thrower is identified with a chip, bags/volume counted, with waste collected at home ('fetch') = **innovation level 3.a**

1 and 10, 1 and 11: the thrower is identified, bags/volume/weight counted, with garbage 'bring' practices = **innovation level 3.b**

1 and 12: evidence of pay-as-you-throw practices = **innovation level 4**

Correspondence table between keyword baskets and innovation levels for the waste collection pilot service

| Combinations of baskets | | | | Innovation level |
|-------------------------|----------|---------|-------|------------------|
| 1 + 11 | 10+11+12 | | | 4 |
| 1+10 | 1+12 | 8+10+11 | | 3b |
| 1+ 7 | 1+8 | 1+ 9 | 7+8+9 | 3a |
| 5 + 6 | | | | 2b |
| 4 + 5 | | | | 2a |
| 1 | | | | 1 |

For Health homecare pilot service

The keywords selected for the Health homecare pilot service were grouped in the following baskets:

1. Related to activities located at home;
2. Related to health care;
3. Care provided at distance;
4. Specialised services provided at a distance;
5. Devices to be used to provide care at a distance;
6. Tools to be self positioned for supporting specialised analyses and controls;
7. Related to interactive, bidirectional, quick and effective communication flows.

Correspondence table between keyword baskets and innovation levels for the Health homecare pilot service

| Combinations of baskets | Innovation level |
|-------------------------|------------------|
| 2 + 4 or 5 + 6 or 7 | 3 |
| 1 + 2 + 5 | 2+ |
| 1 + 2 + 4 | 2 |
| 1 + 2 | 1 |

Annex 4: Example of data for validation and cross- sectional analyses

Example of data for validation and cross-sectional analyses

| Municipality _name | DIFTAR base | Collection method | Volume waste (pp) | Sepa rated (%) | Fetch (freq.) | Sustainability score |
|-----------------------|--|--------------------------------|-------------------------|----------------------|------------------|-------------------------|
| Pekela | number of residents | | | | | |
| Druten | expensive bags and number of residents | minicontain ers | 49 | 0.86 | every 2 weeks | 72 |
| Steenwijkerla nd | volume AND frequency | minicontain ers | 68 | 0.82 | monthly | 70 |
| Bronckhorst | volume AND frequency | minicontain ers | 81 | 0.78 | monthly | 70 |
| Beuningen | expensive bags | minicontain ers AND bags | 90 | 0.77 | every 2 weeks | 69 |
| Heumen | expensive bags and number of residents | bags | 55 | 0.85 | every 2 weeks | 69 |
| Eersel | volume AND frequency | minicontain ers | 52 | 0.82 | monthly | 69 |
| Lochem | volume AND frequency | minicontain ers | 106 | 0.7 | every 2 weeks | 69 |
| Berg en Dal | different systems in parallel | minicontain ers AND bags | 0 | | every 2 weeks | 68 |
| Heeze- Leende | volume AND frequency | | 57 | 0.79 | | 68 |
| Putten | expensive bags | bags | 59 | 0.81 | every week | 68 |
| Zwartewaterl and | volume AND frequency | minicontain ers | 43 | 0.86 | monthly | 68 |
| Olst-Wijhe | volume AND frequency | minicontain ers | 93 | 0.79 | monthly | 68 |
| Beesel | volume AND frequency | minicontain ers | 104 | 0.68 | every 2 weeks | 68 |
| Dalfsen | volume AND frequency | minicontain ers | 57 | 0.86 | monthly | 67 |
| Raalte | volume AND frequency | minicontain ers | 107 | 0.7 | every 3 weeks | 67 |
| Winsum | weight AND frequency | minicontain ers | 0 | | every 2 weeks | 67 |
| Deventer | volume AND frequency | minicontain ers | 80 | 0.74 | every 2 weeks | 66 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|---------------------|--|-------------------------|-------------------|---------------|--|----------------------|
| Beek | weight AND frequency | minicontainers | 108 | 0.68 | every 2 weeks | 66 |
| Twenterand | volume AND frequency | minicontainers | 76 | 0.82 | monthly | 65 |
| Bergeijk | volume AND frequency | minicontainers | 56 | 0.81 | monthly | 65 |
| Brummen | volume AND frequency | minicontainers | 104 | 0.7 | every 2 weeks | 65 |
| Zutphen | volume AND frequency | minicontainers | 130 | 0.62 | every 2 weeks | 65 |
| Ermelo | volume AND frequency | minicontainers | 156 | 0.59 | every 2 weeks | 65 |
| Zuidhorn | weight | minicontainers | 117 | 0.73 | every 2 weeks | 65 |
| Stein | weight AND frequency | minicontainers | 113 | 0.69 | every 2 weeks | 65 |
| Gemert-Bakel | volume AND frequency | minicontainers | 50 | 0.81 | every 2 weeks | 64 |
| Asten | volume AND frequency | | | | | 64 |
| Leek | weight | minicontainers | 118 | 0.68 | every 2 weeks | 64 |
| Lingewaard | volume AND frequency | minicontainers | 122 | 0.67 | every 2 weeks | 64 |
| Grave | expensive bags | bags | 44 | 0.9 | every 2 weeks | 64 |
| Hattem | volume AND frequency | minicontainers | 62 | 0.85 | monthly | 63 |
| Heerde | volume AND frequency | minicontainers | 102 | 0.8 | every 2 weeks | 63 |
| Apeldoorn | volume AND frequency | minicontainers | 132 | 0.65 | every 2 weeks | 63 |
| Tynaarlo | weight | minicontainers | 126 | 0.63 | every 2 weeks | 63 |
| Deurne | volume AND frequency | minicontainers | 77 | 0.7 | monthly | 63 |
| Doesburg | volume AND frequency | minicontainers | 115 | 0.67 | every 3 weeks | 63 |
| Boxmeer | expensive bags | bags | 44 | 0.9 | every 2 weeks | 63 |
| Mill en Sint Hubert | expensive bags and number of residents | bags | 44 | 0.9 | every 2 weeks | 63 |
| Meerssen | expensive bags and number of residents | bags | 94 | 0.82 | every 2 weeks | 62 |
| Cranendonck | volume AND frequency | minicontainers | 75 | 0.81 | monthly | 62 |
| Bladel | volume AND frequency | minicontainers AND bags | 55 | 0.8 | minicontainers monthly, bags every 2 weeks | 62 |
| Marum | weight | minicontainers | 116 | 0.74 | every 2 weeks | 62 |

| Municipality _name | DIFTAR base | Collection method | Volume waste (pp) | Sepa rated (%) | Fetch (freq.) | Sustainability score |
|-----------------------|--|-------------------------|-------------------------|----------------------|---|-------------------------|
| Sittard-Geleen | weight AND frequency | minicontainers AND bags | 127 | 0.63 | minicontainers every 2 weeks, bags weekly | 62 |
| Sint Anthonis | expensive bags | bags | 45 | 0.9 | every 2 weeks | 62 |
| Oost Gelre | volume AND frequency | minicontainers | 43 | 0.86 | monthly | 62 |
| Voorst | volume AND frequency | minicontainers | 112 | 0.75 | every 2 weeks | 62 |
| Hardenberg | volume AND frequency | minicontainers | 108 | 0.71 | every 2 weeks | 62 |
| Buren | volume AND frequency | minicontainers | 118 | 0.72 | every 2 weeks | 62 |
| Boekel | expensive bags | bags | 44 | 0.9 | every 2 weeks | 61 |
| Zundert | expensive bags and number of residents | bags | 53 | 0.86 | every 2 weeks | 61 |
| Nijmegen | expensive bags and number of residents | bags | 92 | 0.69 | every 2 weeks | 61 |
| Zoeterwoude | weight | minicontainers | 82 | 0.71 | every 2 weeks | 61 |
| Staphorst | volume AND frequency | minicontainers | 58 | 0.87 | monthly | 61 |
| Geldermalsen | volume AND frequency | minicontainers | 117 | 0.73 | every 2 weeks | 61 |
| Zaltbommel | volume AND frequency | minicontainers | 117 | 0.72 | every 2 weeks | 61 |
| Voerendaal | volume AND frequency | minicontainers | 127 | 0.71 | every 2 weeks | 61 |
| Culemborg | volume AND frequency | minicontainers | 119 | 0.68 | every 2 weeks | 61 |
| Etten-Leur | volume AND frequency | minicontainers | 111 | 0.75 | every 2 weeks | 60 |
| Nuth | volume AND frequency | minicontainers | 139 | 0.7 | every 2 weeks | 60 |
| Simpelveld | volume AND frequency | minicontainers | 133 | 0.69 | monthly | 60 |
| Zuidplas | volume AND frequency | minicontainers | 96 | 0.68 | every 2 weeks | 60 |
| Aalten | volume AND frequency | minicontainers | 150 | 0.61 | every 2 weeks | 60 |
| Borger-Odoorn | weight | minicontainers | 126 | 0.62 | every 3 weeks | 60 |
| Horst aan de Maas | expensive bags and number of residents | bags | 21 | 0.95 | every 2 weeks | 60 |
| Aalburg | volume AND frequency | minicontainers | 122 | 0.71 | every 2 weeks | 60 |
| Nijkerk | volume AND frequency | minicontainers AND bags | 120 | 0.7 | every 2 weeks | 60 |

| Municipality _name | DIFTAR base | Collection method | Volume waste (pp) | Sepa rated (%) | Fetch (freq.) | Sustainability score |
|-------------------------------------|---|--------------------------|-------------------------|----------------------|------------------|-------------------------|
| Maasdriel | volume AND frequency | minicontain ers | 127 | 0.71 | every 2 weeks | 60 |
| Schinnen | weight AND frequency | minicontain ers | 128 | 0.62 | | 60 |
| Someren | volume AND frequency | minicontain ers | 70 | 0.74 | monthly | 59 |
| Wijchen | volume AND frequency | minicontain ers | 127 | 0.7 | every 2 weeks | 59 |
| Oosterhout | weight AND frequency | minicontain ers | 83 | 0.84 | every 2 weeks | 59 |
| Ommen | volume AND frequency | minicontain ers | 63 | 0.87 | every 2 weeks | 59 |
| West Maas en Waal | volume AND frequency | minicontain ers | 113 | 0.76 | every 2 weeks | 59 |
| Geldrop- Mierlo | volume AND frequency | minicontain ers | 112 | 0.68 | every 2 weeks | 59 |
| Vaals | volume, frequency and number of residents | minicontain ers | 138 | 0.67 | every 2 weeks | 59 |
| Mook en Middelaar | weight | minicontain ers | 111 | 0.7 | every 2 weeks | 59 |
| Woensdrecht | volume AND frequency | | | | | 59 |
| Venlo | number of residents | duo(split)co ntainers | 154 | 0.64 | every 2 weeks | 58 |
| Nuenen, Gerwen en Nederwetten | volume AND frequency | minicontain ers | 78 | 0.78 | monthly | 58 |
| Lingewaal | volume AND frequency | minicontain ers | 126 | 0.7 | every 2 weeks | 58 |
| Kampen | volume AND frequency | minicontain ers | 109 | 0.68 | every 2 weeks | 58 |
| Grootevast | weight | | | | | 58 |
| Waalre | number of residents | minicontain ers | 120 | 0.74 | monthly | 58 |
| Werkendam | volume AND frequency | minicontain ers | 110 | 0.76 | every 2 weeks | 58 |
| Haren | weight | minicontain ers | 151 | 0.56 | every 2 weeks | 58 |
| Bernheze | weight AND frequency | minicontain ers | 119 | 0.67 | every 2 weeks | 58 |
| Maasgouw | volume AND frequency | minicontain ers | 139 | 0.7 | every 2 weeks | 58 |
| Cuijk | expensive bags | bags | 44 | 0.9 | every 2 weeks | 58 |
| Son en Breugel | number of residents | minicontain ers | 105 | 0.8 | monthly | 57 |
| Leudal | volume AND frequency | minicontain ers | 116 | 0.72 | every 2 weeks | 57 |
| Hengelo | volume AND frequency | minicontain ers | 128 | 0.7 | every 2 weeks | 57 |
| Peel en Maas | number of residents | minicontain ers | 97 | 0.79 | every 2 weeks | 57 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|---------------------|--------------------------------|-------------------------|-------------------|---------------|---|----------------------|
| Baarle-Nassau | volume | minicontainers | 152 | 0.7 | every 2 weeks | 57 |
| Neder-Betuwe | volume AND frequency | minicontainers | 120 | 0.71 | every 2 weeks | 57 |
| Onderbanken | volume AND frequency | minicontainers | 151 | 0.67 | every 2 weeks | 57 |
| Roerdalen | volume AND frequency | minicontainers | 142 | 0.67 | monthly | 57 |
| Oirschot | volume AND frequency | minicontainers | 62 | 0.8 | monthly | 57 |
| Montferland | volume AND frequency | duo(split)containers | 145 | 0.68 | every week | 57 |
| Meppel | fixed tariff | minicontainers AND bags | 139 | 0.66 | Minicontainers monthly, bags every week | 56 |
| Sint-Michielsgestel | volume AND frequency | minicontainers | 128 | 0.64 | every 2 weeks | 56 |
| Nederweert | weight | duo(split)containers | 116 | 0.69 | every week | 56 |
| Eijsden-Margraten | weight and number of residents | minicontainers | 93 | 0.78 | every 2 weeks | 56 |
| Waalwijk | weight AND frequency | minicontainers | 97 | 0.72 | monthly | 56 |
| Veere | number of residents | minicontainers | 171 | 0.71 | every 2 weeks | 56 |
| Woudrichem | volume AND frequency | minicontainers | 126 | 0.68 | every 2 weeks | 56 |
| Gulpen-Wittem | number of residents | duo(split)containers | 0 | | every 2 weeks | 56 |
| Tubbergen | volume AND frequency | minicontainers | 70 | 0.78 | every week | 56 |
| Tiel | volume AND frequency | minicontainers | 128 | 0.67 | every 2 weeks | 56 |
| Noordwijk | number of residents | minicontainers | 196 | 0,49 | every 2 weeks | 55 |
| De Fryske Marren | volume | minicontainers | 163 | 0.68 | every 2 weeks | 55 |
| Oegstgeest | volume | minicontainers | 181 | 0.55 | every 2 weeks | 55 |
| Haaren | volume AND frequency | minicontainers | 99 | 0.76 | every 2 weeks | 55 |
| Brunssum | volume AND frequency | minicontainers | 155 | 0.64 | every 2 weeks | 55 |
| Hilvarenbeek | volume AND frequency | minicontainers | 107 | 0.76 | every 3 weeks | 55 |
| Scherpenzeel | number of residents | minicontainers | 178 | 0.58 | every 2 weeks | 54 |
| Wierden | volume | minicontainers | 173 | 0.63 | every 2 weeks | 54 |
| Midden-Delfland | number of residents | | 179 | 0.53 | | 54 |
| Reusel-De Mierden | volume AND frequency | bags | 19 | 0.89 | every 2 weeks | 54 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|-------------------|--|----------------------|-------------------|---------------|---------------|----------------------|
| Teylingen | volume | minicontainers | 165 | 0.62 | every 2 weeks | 54 |
| Rheden | volume | minicontainers | 178 | 0.59 | every 2 weeks | 53 |
| Rucphen | volume AND frequency | minicontainers | 155 | 0.64 | every 2 weeks | 53 |
| Heerlen | volume, frequency and number of residents | minicontainers | 163 | 0.61 | every 2 weeks | 53 |
| Leeuwarderadeel | number of residents | minicontainers | 196 | 0.62 | every 2 weeks | 53 |
| Lisse | number of residents | minicontainers | 156 | 0.62 | every 2 weeks | 53 |
| Aalsmeer | volume | minicontainers | 180 | 0.58 | every 2 weeks | 53 |
| Boxtel | volume, frequentie and number of residents | | 141 | 0.64 | | 53 |
| Smallingerland | number of residents | minicontainers | 165 | 0.71 | every 2 weeks | 52 |
| Wageningen | number of residents | minicontainers | 126 | 0.64 | every 2 weeks | 52 |
| Katwijk | number of residents | minicontainers | 192 | 0.49 | every 2 weeks | 52 |
| Heusden | volume AND frequency | minicontainers | 154 | 0.66 | every 2 weeks | 52 |
| Heiloo | number of residents | minicontainers | 180 | 0.62 | every 2 weeks | 52 |
| Ouder-Amstel | number of residents | | 259 | 0.43 | | 52 |
| Ameland | number of residents | minicontainers | 0 | | every 2 weeks | 52 |
| Ten Boer | volume | minicontainers | 224 | 0.53 | every 2 weeks | 52 |
| Overbetuwe | volume AND frequency | minicontainers | 130 | 0.7 | every 2 weeks | 52 |
| Bunnik | number of residents | minicontainers | 185 | 0.64 | every 2 weeks | 52 |
| Leusden | number of residents | bags | 144 | 0.69 | every week | 51 |
| Loon op Zand | number of residents | minicontainers | 167 | 0.65 | every 2 weeks | 51 |
| Renkum | number of residents | duo(split)containers | 180 | 0.6 | every week | 51 |
| Weststellingwerf | number of residents | minicontainers | 252 | 0.54 | every 2 weeks | 51 |
| Groningen | number of residents | minicontainers | 216 | 0.39 | every 2 weeks | 51 |
| Hof van Twente | volume | minicontainers | 178 | 0.64 | every 2 weeks | 51 |
| Gilze en Rijen | volume AND frequency | | 120 | 0.75 | | 51 |
| Opmeer | number of residents | minicontainers | 125 | 0.74 | every 2 weeks | 51 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|--------------------------------|--|----------------------|-------------------|---------------|---------------|----------------------|
| Landerd | number of residents | minicontainers | 213 | 0.62 | every 2 weeks | 51 |
| Hillegom | number of residents | minicontainers | 189 | 0.56 | every 2 weeks | 51 |
| Oldebroek | number of residents | minicontainers | 219 | 0.55 | every 2 weeks | 51 |
| Harderwijk | number of residents | minicontainers | 207 | 0.53 | every 2 weeks | 51 |
| Epe | number of residents | minicontainers | 188 | 0.64 | every 3 weeks | 51 |
| Soest | number of residents | minicontainers | 189 | 0.61 | every 2 weeks | 51 |
| Houten | number of residents | minicontainers | 184 | 0.61 | every 2 weeks | 50 |
| De Ronde Venen | number of residents | minicontainers | 182 | 0.6 | every 2 weeks | 50 |
| Elburg | number of residents | minicontainers | 229 | 0.51 | every 2 weeks | 50 |
| Opsterland | volume AND frequency | minicontainers | 163 | 0.68 | every 2 weeks | 50 |
| Kollumerland en Nieuwkruisland | number of residents | minicontainers | 164 | 0.71 | every 2 weeks | 50 |
| Krimpenerwaard | different systems in parallel | | 165 | 0.63 | | 49 |
| Dongen | number of residents | bags | 141 | 0.72 | every week | 49 |
| Castricum | number of residents | minicontainers | 202 | 0.59 | every 2 weeks | 49 |
| Roermond | fixed tariff | duo(split)containers | 189 | 0.57 | every week | 49 |
| Oudewater | number of residents | minicontainers | 199 | 0.57 | every 2 weeks | 49 |
| Lopik | volume | minicontainers | 163 | 0.69 | every 2 weeks | 49 |
| De Marne | volume AND frequency | minicontainers | 219 | 0.55 | every 2 weeks | 49 |
| Bloemendaal | volume | minicontainers | 188 | 0.56 | every 2 weeks | 49 |
| Kapelle | volume, frequentie and number of residents | minicontainers | 126 | 0.7 | every 2 weeks | 49 |
| Almere | fixed tariff | duo(split)containers | 179 | 0.58 | every 2 weeks | 48 |
| Utrechtse Heuvelrug | number of residents | minicontainers | 125 | 0.73 | every 2 weeks | 48 |
| Veldhoven | number of residents | minicontainers | 112 | 0.73 | monthly | 48 |
| Zeewolde | number of residents | minicontainers | 150 | 0.68 | every 2 weeks | 48 |
| Woerden | number of residents | minicontainers | 178 | 0.6 | every 2 weeks | 48 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|--------------------|----------------------|-------------------------|-------------------|---------------|---|----------------------|
| Voorschoten | number of residents | minicontainers AND bags | 244 | 0.39 | minicontainers every 2 weeks, bags weekly | 48 |
| Baarn | fixed tariff | minicontainers | 196 | 0.57 | every 2 weeks | 48 |
| Eemnes | number of residents | minicontainers | 193 | 0.6 | every 2 weeks | 48 |
| Noordenveld | number of residents | minicontainers | 217 | 0.59 | every 2 weeks | 48 |
| Hilversum | number of residents | minicontainers AND bags | 199 | 0.56 | minicontainers every 2 weeks, bags weekly | 48 |
| Huizen | number of residents | minicontainers AND bags | 199 | 0.56 | minicontainers every 2 weeks, bags weekly | 48 |
| Goirle | number of residents | minicontainers | 0 | | every 2 weeks | 48 |
| Helmond | volume | minicontainers | 191 | 0.54 | every 2 weeks | 48 |
| Bergen op Zoom | volume AND frequency | minicontainers | 170 | 0.65 | every 2 weeks | 48 |
| Blaricum | number of residents | minicontainers AND bags | 199 | 0.56 | minicontainers every 2 weeks, bags weekly | 48 |
| Laren | number of residents | minicontainers AND bags | 199 | 0.56 | minicontainers every 2 weeks, bags weekly | 48 |
| Weesp | number of residents | minicontainers AND bags | 199 | 0.56 | minicontainers every 2 weeks, bags weekly | 48 |
| Wijdmeren | number of residents | minicontainers AND bags | 199 | 0.56 | minicontainers every 2 weeks, bags weekly | 48 |
| Bergen op Zoom | volume AND frequency | minicontainers | 170 | 0.65 | every 2 weeks | 48 |
| Heemstede | volume | minicontainers AND bags | 195 | 0.6 | minicontainers every 2 weeks, bags weekly | 47 |
| Dongeradeel | number of residents | minicontainers | 184 | 0.62 | every 2 weeks | 47 |
| Schouwen-Duiveland | volume | minicontainers | 245 | 0.61 | every 2 weeks | 47 |
| Ooststellingwerf | volume AND frequency | minicontainers | 173 | 0.65 | every 2 weeks | 47 |
| Halderberge | volume AND frequency | minicontainers | | | every 2 weeks | 47 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|-------------------|----------------------|-------------------------|-------------------|---------------|---|----------------------|
| Dronten | fixed tariff | minicontainers | 169 | 0.62 | every 2 weeks | 47 |
| Rijssen-Holten | number of residents | minicontainers | 194 | 0.68 | every 2 weeks | 46 |
| Assen | number of residents | minicontainers | 169 | 0.59 | every 2 weeks | 46 |
| Barneveld | number of residents | minicontainers | 191 | 0.59 | every 2 weeks | 46 |
| Uden | number of residents | minicontainers | 207 | 0.55 | every 2 weeks | 46 |
| Amstelveen | number of residents | minicontainers | 201 | 0.49 | every 2 weeks | 46 |
| Hellendoorn | volume | minicontainers | 181 | 0.64 | every 2 weeks | 46 |
| De Wolden | number of residents | minicontainers | 172 | 0.73 | every 2 weeks | 46 |
| Leeuwarden | number of residents | minicontainers | 214 | 0.51 | every 2 weeks | 46 |
| Best | fixed tariff | minicontainers | 191 | 0.59 | every 2 weeks | 45 |
| Oldenzaal | number of residents | minicontainers AND bags | 185 | 0.62 | minicontainers every 2 weeks, bags weekly | 45 |
| Ede | number of residents | minicontainers | 196 | 0.55 | every 2 weeks | 45 |
| Stichtse Vecht | number of residents | minicontainers | 214 | 0.45 | every 2 weeks | 45 |
| Koggenland | number of residents | minicontainers | 138 | 0.7 | every 2 weeks | 45 |
| Achtkarspeelen | number of residents | minicontainers | 207 | 0.68 | every 2 weeks | 45 |
| Zeist | number of residents | minicontainers | 203 | 0.56 | every 2 weeks | 45 |
| Haaksbergen | volume | minicontainers | 216 | 0.58 | every 2 weeks | 45 |
| Oldambt | volume AND frequency | minicontainers | 146 | 0.72 | every 2 weeks | 45 |
| Borne | fixed tariff | minicontainers AND bags | 222 | 0.61 | minicontainers every 2 weeks, bags weekly | 44 |
| Tilburg | fixed tariff | duo(split)containers | 156 | 0.58 | every week | 44 |
| Waddinxveen | number of residents | minicontainers AND bags | 209 | 0.47 | every 2 weeks | 44 |
| Dinkelland | volume AND frequency | minicontainers | 66 | 0.79 | every 2 weeks | 44 |
| Alphen-Chaam | number of residents | minicontainers | 209 | 0.59 | every 2 weeks | 44 |
| Beemster | number of residents | duo(split)containers | 232 | 0.47 | every 2 weeks | 44 |
| Delft | number of residents | | 210 | 0.39 | | 44 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|--------------------|----------------------|-------------------------|-------------------|---------------|---|----------------------|
| Berkelland | volume | minicontainers | 252 | 0.52 | every 2 weeks | 44 |
| Cromstrijen | volume AND frequency | minicontainers | 135 | 0.72 | every 2 weeks | 44 |
| Binnenmaas | volume AND frequency | minicontainers | 133 | 0.71 | every 2 weeks | 44 |
| Aa en Hunze | number of residents | minicontainers | 162 | 0.72 | monthly | 43 |
| Langedijk | number of residents | minicontainers | 205 | 0.6 | every 2 weeks | 43 |
| Delfzijl | number of residents | minicontainers | 251 | 0.55 | every 2 weeks | 43 |
| Geertruidenberg | number of residents | minicontainers | 199 | 0.55 | every 2 weeks | 43 |
| Goeree-Overflakkee | number of residents | minicontainers | 218 | 0.55 | every 2 weeks | 43 |
| Gouda | number of residents | minicontainers AND bags | 199 | 0.48 | minicontainers every 2 weeks, bags weekly | 43 |
| Urk | number of residents | minicontainers | 209 | 0.46 | every 2 weeks | 43 |
| Veenendaal | volume AND frequency | minicontainers | 91 | 0.72 | every 2 weeks | 43 |
| Schagen | number of residents | minicontainers | 170 | 0.66 | every 2 weeks | 43 |
| Oude IJsselstreek | number of residents | minicontainers | 248 | 0.53 | every 2 weeks | 43 |
| Alkmaar | number of residents | | 207 | 0.45 | | 43 |
| Appingedam | number of residents | minicontainers | 0 | | every 2 weeks | 43 |
| Oss | volume | minicontainers | 167 | 0.67 | every 3 weeks | 43 |
| Medemblik | number of residents | minicontainers | 172 | 0.66 | every 2 weeks | 43 |
| Enschede | number of residents | minicontainers | 207 | 0.47 | every 2 weeks | 42 |
| Vianen | volume | minicontainers | 217 | 0.56 | every 2 weeks | 42 |
| Bedum | fixed tariff | minicontainers | 256 | 0.52 | every 2 weeks | 42 |
| Hulst | number of residents | minicontainers | 235 | 0.66 | every 2 weeks | 42 |
| Waterland | number of residents | minicontainers AND bags | 217 | 0.56 | minicontainers every 2 weeks, bags weekly | 42 |
| Westland | number of residents | minicontainers | 208 | 0.54 | every 2 weeks | 42 |
| Veendam | number of residents | minicontainers | 248 | 0.53 | every 2 weeks | 42 |
| Weert | number of residents | duo(split)containers | 0 | | every week | 42 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|--------------------|----------------------|-------------------------|-------------------|---------------|---|----------------------|
| Borsele | number of residents | minicontainers | 202 | 0.67 | every 2 weeks | 42 |
| Hoogeveen | number of residents | minicontainers | 194 | 0.61 | every 2 weeks | 42 |
| Wijk bij Duurstede | fixed tariff | minicontainers | 141 | 0.67 | every 3 weeks | 41 |
| Goes | number of residents | minicontainers | 208 | 0.61 | every 2 weeks | 41 |
| Zwolle | number of residents | minicontainers | 193 | 0.51 | every 2 weeks | 41 |
| Eindhoven | number of residents | minicontainers AND bags | 200 | 0.5 | minicontainers every 2 weeks, bags weekly | 41 |
| Nieuwegein | number of residents | minicontainers | 223 | 0.44 | every 2 weeks | 41 |
| Korendijk | volume AND frequency | minicontainers | 134 | 0.71 | every 2 weeks | 41 |
| Doetinchem | number of residents | | 164 | 0.64 | | 41 |
| Strijen | volume AND frequency | minicontainers | 133 | 0.71 | every 2 weeks | 41 |
| Westerveld | number of residents | minicontainers | 156 | 0.71 | monthly | 40 |
| Hoorn | number of residents | minicontainers | 183 | 0.6 | every 2 weeks | 40 |
| Heemskerk | number of residents | minicontainers | 230 | 0.5 | every 2 weeks | 40 |
| Noord-Beveland | number of residents | minicontainers | 264 | 0.63 | every 2 weeks | 40 |
| Bergen (NH.) | number of residents | minicontainers | 210 | 0.61 | every 2 weeks | 40 |
| Sluis | number of residents | minicontainers | 309 | 0.61 | every 2 weeks | 40 |
| Bunschoten | number of residents | minicontainers | 183 | 0.59 | every 2 weeks | 40 |
| Zoetermeer | number of residents | minicontainers | 222 | 0.39 | every 2 weeks | 39 |
| Uithoorn | fixed tariff | minicontainers | 241 | 0.49 | every 2 weeks | 39 |
| Reimerswaal | number of residents | minicontainers | 215 | 0.67 | every 2 weeks | 39 |
| Ferwerderadiel | number of residents | minicontainers | 176 | 0.65 | every 2 weeks | 39 |
| Zederik | number of residents | minicontainers | 182 | 0.64 | every 2 weeks | 39 |
| Sliedrecht | number of residents | | 196 | 0.58 | | 39 |
| Uitgeest | number of residents | minicontainers | 204 | 0.55 | every 2 weeks | 39 |
| Heerhugowaard | number of residents | minicontainers | 216 | 0.54 | every 2 weeks | 39 |
| Valkenswaard | number of residents | minicontainers | 0 | | every 2 weeks | 39 |
| Oud-Beijerland | volume AND frequency | minicontainers | 133 | 0.71 | every 2 weeks | 39 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|------------------------|----------------------|-------------------------|-------------------|---------------|---|----------------------|
| Moerdijk | number of residents | duo(split)containers | 219 | 0.58 | every week | 38 |
| Lelystad | number of residents | minicontainers | 209 | 0.57 | monthly | 38 |
| Alphen aan den Rijn | number of residents | minicontainers AND bags | | | minicontainers every 2 weeks, bags weekly | 38 |
| Winterswijk | volume AND frequency | minicontainers | 68 | 0.83 | every 2 weeks | 38 |
| Hardinxveld-Giessendam | fixed tariff | minicontainers | 181 | 0.63 | every 2 weeks | 38 |
| Eemsmond | number of residents | minicontainers | 246 | 0.55 | every week | 38 |
| Harlingen | number of residents | minicontainers | 228 | 0.53 | every 2 weeks | 38 |
| Wassenaar | number of residents | | 282 | 0.5 | | 38 |
| Coevorden | number of residents | minicontainers | 254 | 0.66 | every 2 weeks | 38 |
| Lansingerland | number of residents | | 202 | 0.51 | | 37 |
| Arnhem | number of residents | minicontainers AND bags | 208 | 0.47 | minicontainers every 2 weeks, bags weekly | 37 |
| 's-Hertogenbosch | number of residents | minicontainers | 225 | 0.46 | every 2 weeks | 37 |
| Leerdam | fixed tariff | minicontainers | 188 | 0.62 | every 2 weeks | 37 |
| Edam-Volendam | number of residents | minicontainers AND bags | 252 | 0.57 | minicontainers every 2 weeks, bags weekly | 37 |
| Amersfoort | number of residents | minicontainers | 213 | 0.52 | every 2 weeks | 37 |
| Terschelling | fixed tariff | minicontainers | 355 | 0.43 | every 2 weeks | 36 |
| Terneuzen | number of residents | minicontainers | 232 | 0.59 | every 2 weeks | 36 |
| Schiermonnikoog | number of residents | duo(split)containers | 431 | 0.57 | every 2 weeks | 36 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|--------------------------------|---------------------|-------------------------|-------------------|---------------|---|----------------------|
| Renswoude | number of residents | minicontainers | 248 | 0.51 | every 2 weeks | 36 |
| Utrecht | number of residents | minicontainers AND bags | 216 | 0.41 | every week | 35 |
| Gorinchem | number of residents | minicontainers | 205 | 0.62 | every 2 weeks | 35 |
| Zandvoort | number of residents | minicontainers AND bags | 264 | 0.35 | minicontainers every 2 weeks, bags weekly | 35 |
| Landsmeer | number of residents | minicontainers | 0 | | every 2 weeks | 35 |
| Zwijndrecht | number of residents | | | | | 35 |
| Alblasserdam | number of residents | minicontainers | 219 | 0.53 | every 2 weeks | 34 |
| Haarlem | number of residents | minicontainers AND bags | 260 | 0.33 | minicontainers every 2 weeks, bags weekly | 34 |
| Diemen | number of residents | | 232 | 0.32 | | 34 |
| Hellevoetsluis | fixed tariff | minicontainers | 232 | 0.51 | every 2 weeks | 34 |
| Leiden | number of residents | minicontainers AND bags | 242 | 0.33 | Minicontainers every 2 weeks, bags meerder keren per week | 34 |
| Brielle | fixed tariff | minicontainers | 261 | 0.48 | every 2 weeks | 33 |
| Tholen | number of residents | | | | | 33 |
| Zaanstad | fixed tariff | minicontainers | 242 | 0.43 | every 2 weeks | 33 |
| Papendrecht | fixed tariff | minicontainers | 233 | 0.49 | every 2 weeks | 32 |
| Haarlemmerliede en Spaarnwoude | number of residents | Mini containers | 300 | 0.36 | | 32 |

| Municipality_name | DIFTAR base | Collection method | Volume waste (pp) | Separated (%) | Fetch (freq.) | Sustainability score |
|------------------------|---------------------|-------------------------|-------------------|---------------|---------------|----------------------|
| Velsen | fixed tariff | Mini containers | 0 | | every 2 weeks | 31 |
| Beverwijk | number of residents | minicontainers | 287 | 0.42 | every 2 weeks | 31 |
| Schiedam | number of residents | minicontainers | 264 | 0.3 | every 2 weeks | 31 |
| Amsterdam | number of residents | | 256 | 0.18 | | 30 |
| Vlissingen | fixed tariff | minicontainers | 285 | 0.48 | every 2 weeks | 28 |
| Den Helder | number of residents | minicontainers | 270 | 0.43 | every 2 weeks | 28 |
| Maassluis | number of residents | minicontainers | 263 | 0.41 | every 2 weeks | 28 |
| Vlaardingen | number of residents | minicontainers | 0 | | every 2 weeks | 25 |
| Capelle aan den IJssel | number of residents | | | | | 25 |
| Purmerend | fixed tariff | minicontainers | 220 | 0.53 | every week | 23 |
| Rotterdam | number of residents | minicontainers AND bags | 303 | 0.21 | | 23 |
| Hendrik-Ido-Ambacht | number of residents | bags | 243 | 0.49 | every week | 22 |
| Krimpen aan den IJssel | number of residents | minicontainers | 286 | 0.46 | every 2 weeks | 21 |
| Nissewaard | fixed tariff | minicontainers | 240 | 0.39 | every 2 weeks | 21 |

Annex 5: Calculation of regional and national averages of innovativeness levels

Innovativeness levels, which values ranged from 1 = undifferentiated waste collection to 4 = pay-as-you-throw, with 0 as not visited website or not found keyword, were available for each surveyed municipality. Since the number and size of municipalities with available data varies by region, regional averages were computed using the following formula.

Let:

I_{kr} be the innovativeness level in municipality k of region r

n_{kr} the population size in municipality k of region r

n_r the total population size of all the participating municipalities in region r

n the total population size of all participating municipalities

In this notation, it is defined by k the index of the municipality and with the r that of the region, with the total number of municipalities varying by region, i.e. the municipality index in region r varies from 1 to K_r ($k=1,2,\dots, K_r$), and the region index varies from 1 to 20 ($r=1,2,\dots,20$).

Then the regional average for region r , \bar{I}_r , was calculated as the weighted average the municipality-specific innovativeness levels I_{kr} , with weights given by the municipalities' sizes relative to the regional size, i.e. n_{kr} / n_r , summed over the total number of municipalities in the region. Formally:

$$\bar{I}_r = \sum_{k=1}^{K_r} I_{kr} (n_{kr}/n_r)$$

The national average was then calculated as the average of the regional-specific innovativeness levels, with weights given by the region relative sizes (i.e. n_r/n). Formally:

$$\bar{I} = \sum_{r=1}^{20} \bar{I}_r (n_r/n).$$

Getting in touch with the EU

In person

All over the European Union there are hundreds of Europe Direct Information Centres. You can find the address of the centre nearest you at: <http://europa.eu/contact>

On the phone or by e-mail

Europe Direct is a service that answers your questions about the European Union. You can contact this service

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696 or
- by electronic mail via: <http://europa.eu/contact>

Finding information about the EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: <http://europa.eu>

EU Publications

You can download or order free and priced EU publications from EU Bookshop at: <http://bookshop.europa.eu>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see <http://europa.eu/contact>)

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: <http://eur-lex.europa.eu>

Open data from the EU

The EU Open Data Portal (<http://data.europa.eu/euodp/en/data>) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.

Measuring public innovation in the EU: the STARPIN methodology

This work accounts for, and synthesises the main results of, the StarPIN project — Statistical Reporting on Public Innovation — developed by the Dipartimento di Economia, Società, Politica (DESP) of the University of Urbino, Italy, with the support of dialogic — innovatie. interactie, Utrecht, the Netherlands. The StarPIN project was supported by Eurostat's Unit working on innovation statistics. The StarPIN team was composed of Annaflavia Bianchi, Claudio Cozza, Giovanni Marin, Robbin Te Velde, Antonello Zanfei, and Emy Zecca. Gregor Kyi and Giulio Perani were the Eurostat project managers and Stefania Panaitescu ensured editing and support.

The project proposes a theoretical and a methodological framework for enriching the measurement ability of public sector innovation, and discusses the implications for data collection and analysis. A preliminary test of the framework is carried out by means of pilot applications to specific public services.

The project focuses on public service innovation as a key locus of value creation for society, choosing an object-based rather than the more commonly adopted subject-based approach. Issues concerning the measurement of public service innovation are addressed in statistical terms and consistently with the official public sector functions classifications. The approach places particular emphasis on web-scraping to capture the technological level of websites used by public administrations, and to evaluate the degree of innovativeness of specific services within selected public functions. Data collected through web-scraping can be combined with administrative data at the level of individual services. The paper presents pilot applications to specific public services in a limited number of countries.

The paper is structured according to the following steps: First a theoretical framework is sketched for the analysis of innovation in the public sector, focusing on service innovation and public value creation. Second, the consistency between public functions, public institutions and public service classifications is discussed. Third, the indicators of innovativeness are generated for selected public services. Fourth, the paper illustrates the main lessons learnt from the pilot application of the methodological framework, and draws implications for future research. A user manual is annexed to the paper for replicating the pilot tests described in the paper.

For more information

<https://ec.europa.eu/eurostat/>