



ESPOO
ESBO



Producing and utilising customer information: a handbook





#MakeWithEspoo collects the results of the experiments conducted in Espoo in order to create tools of development, management and consulting according to the City as a Service objective. It is an umbrella of the co-creation of frameworks, handbooks and tools created in 6Aika projects that reforms activities in a socially, culturally, ecologically and economically sustainable way.

6Aika is a joint strategy of the six biggest cities in Finland - Helsinki, Espoo, Tampere, Vantaa, Oulu and Turku - aiming at the development of more open and intelligent services. The aim is to create new competence, business and jobs in Finland. At the end of 2017, there were over 30 ongoing 6Aika projects and their total budget was over 57 million euros.



Foreword

In the next decades, the world will change more than it has changed in the last centuries. Inspiring innovations change established operations, creating new opportunities for building a sustainable and human-oriented future.

The public sector must actively search for a new role in the changes of individuals, communities, businesses, society and the environment. In the place of traditional, siloed and administration-based activity, we need new, open and effective ways to understand customer relationships, information and activity. Open activity strengthens participation, responsibility and trust. It clears the way for customer-relationship-based services, new business operations and the new role of public operations.

In Espoo, we create new success stories by introducing different actors, operations and networks to each other. The MakeWithEspoo methods and tools can help us to ensure that we do the right things, with the right people, at the right time and using the right skills.

The aim of the handbook on the production and use of customer information is to make customer relationships based on customers and not the organisation. In service development at the municipal level, the customers, their well-being and needs are at the centre, not the organisational structure. In this way, it is possible to produce services both proactively and economically. We have wanted to develop a solution that can be used to systematically collect and analyse information on customers and that can be utilised by both the city and various interest groups, for instance in the development of multichannel services or in creating new business.

My warmest thanks to all who participated in the development, and especially to Data Analytics Consultant Tomas Lehtinen, Project Director Marke Kaukonen, Development Manager Arja Kaikkonen and our cooperation partners KPMG Oy and Tieto Oyj.

Päivi Sutinen

PhD, EMBA, Service Development Director
City of Espoo

ISBN [number here] (print)
ISBN [number here] (PDF)

This work is licensed under the Creative Commons BY 4.0 International licence.



6Aika



Vipuvoimaa
EU:lta
2014–2020



Contents

1 Background and objectives	5
2 Producing and refining information	6
2.1 The operational model for refining data	8
2.2 The process of refining data	9
2.3 Cases of the use of data; users of the data	14
2.4 Resources and competences	20
2.5 The technology of refining data	23
2.6 Data protection and privacy in data refining	23
3 Making use of customer information	35
3.1 Management of customer relationships	35
3.1.1 Customer information as a part of customer management	37
3.1.2 Grouping of customers	38
3.1.3 Prediction of customer needs	45
3.1.4 Monitoring of customer relationships	46
3.2 Arranging services	47
3.2.1 Planning and monitoring of services	48
3.2.2 Service development	49
3.2.3 Commodification and tailoring of services	54
3.2.4 Managing the development of activity	55
3.3 Service production	56
3.3.1 Utilising customer information	56
3.3.2 Service levels	56
3.3.3 Management and monitoring of service production	57
Appendix 1: Benefit, risk and cost estimate form for analysed data	60

1



Background and objectives

As part of the 6Aika Open Participation and Customership spearhead project, the Service Development Unit of the Mayor's Office of the City of Espoo launched the project "A framework for developing services and the management of customer relationships with data" in April 2017. The project was implemented in cooperation with experts from KPMG Oy. The project has been supported by three sub-projects: the demographic and business analyses conducted with Fonecta, the development of customer-oriented services by the Espoo City Museum as well as the artificial intelligence and data pool of Tieto Oyj.

A framework for developing services and the management of customer relationships with data

The framework for developing services and the management of customer relationships with data describes an operational model for refining and utilising customer information in the development of both customer relationships and services. The framework describes the areas of customer and service development that can make use of customer data, the process of refining data into knowledge and the competences required to develop activity in a data-driven way.

The preparation of the framework has utilised available tools and methods as applicable, including the Kunnan johtamisen viitearkkitehtuuri (Reference architecture for municipal management, JVA).

Representatives of municipal branches, businesses and different interest groups in Espoo have been interviewed, and business workshops have also been arranged to assist in the preparation of the framework.

Two separate reports have been prepared on the framework:

Framework for customer information knowledge management which aims to help to identify uses for management by data. Management by data can help you to understand customer needs better, develop better services and create new vitality for citizens, businesses and the third sector.

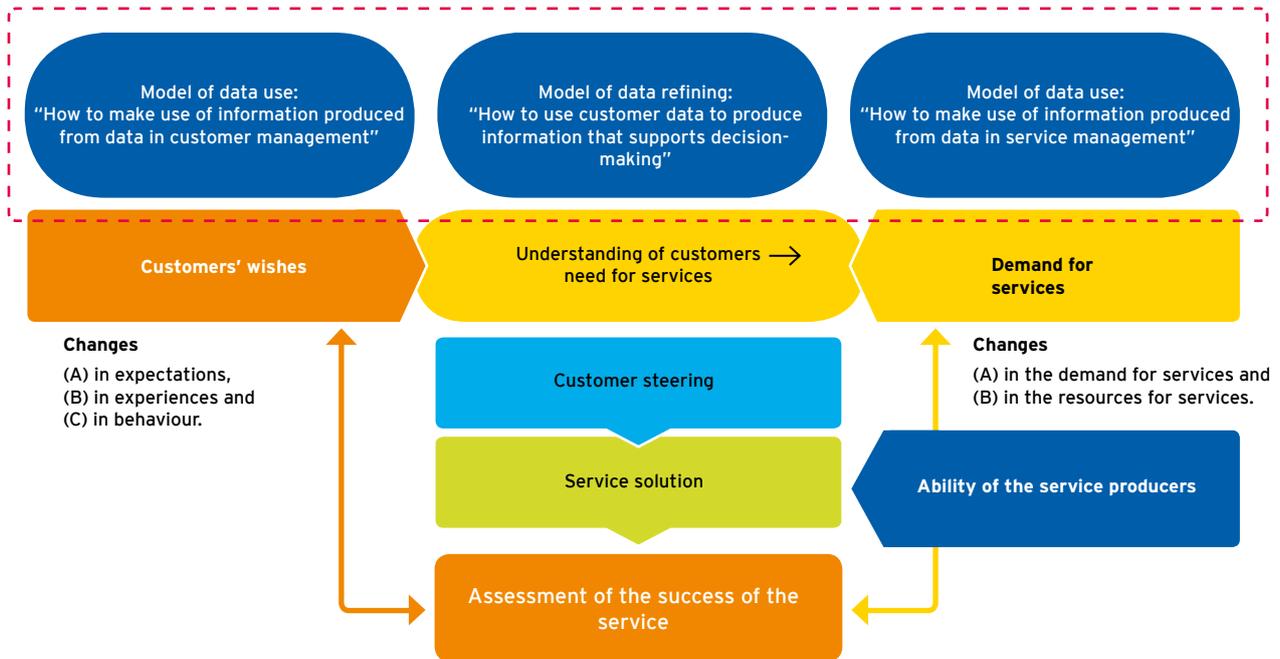
Producing and utilising customer information: a handbook, which aims to function as a handbook and guide to refining data, help with the systematic implementation of the refining process as well as give guidance on conformity with privacy and data security. As an appendix, the handbook also contains a benefit, risk and cost estimate form for assessing new objects of analysis.

2

.....

Producing and refining information

The production and utilisation of customer information (management by data) requires both refining data and making use of the refined data in the processes and practices of the organisation.



The refining of data aims at producing a better understanding of customers that can be utilised in developing both customer relationships and services.

The section on producing and refining information describes the competences required in developing activity based on data. The competences include operational models, processes, tools, resources and skills that can be used to retrieve, combine and analyse data for different purposes and user groups.

The section also describes the different models of administering the production of information and the competences and resources required for the production of information.

In addition, an operational model for processing customer data is created in the section. It supports the refining of data with data security and privacy as a starting point.

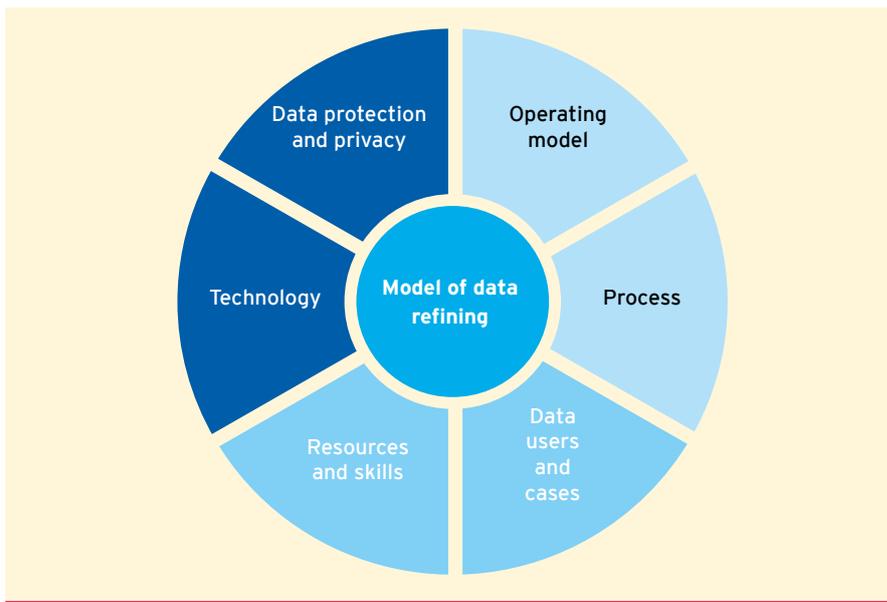
During the project, the processes of the section have been tested in a practical pilot project that used the analysis of data from multiple registers in order to identify the customer paths of the customers of child protective services in Espoo and understand how the native language of the customers affects the customer path.

Refining data means that the data required can be gathered. Its quality in the source systems must also be sufficiently high. The quality of information strongly affects the results and findings produced through analyses. This handbook does not describe the management of data related to the availability of information or to quality control.

This section and the model presented therein are divided into six sub-areas. They describe the competences needed when an activity is developed on the basis of data.

The areas are the following:

- Function or organisation model
- Functional process of data refining
- Making use of data: users and cases
- The resources and skills needed to refine data
- The technologies needed to refine data
- Taking data protection and privacy into account when refining data.



The model of refining data and the competences required

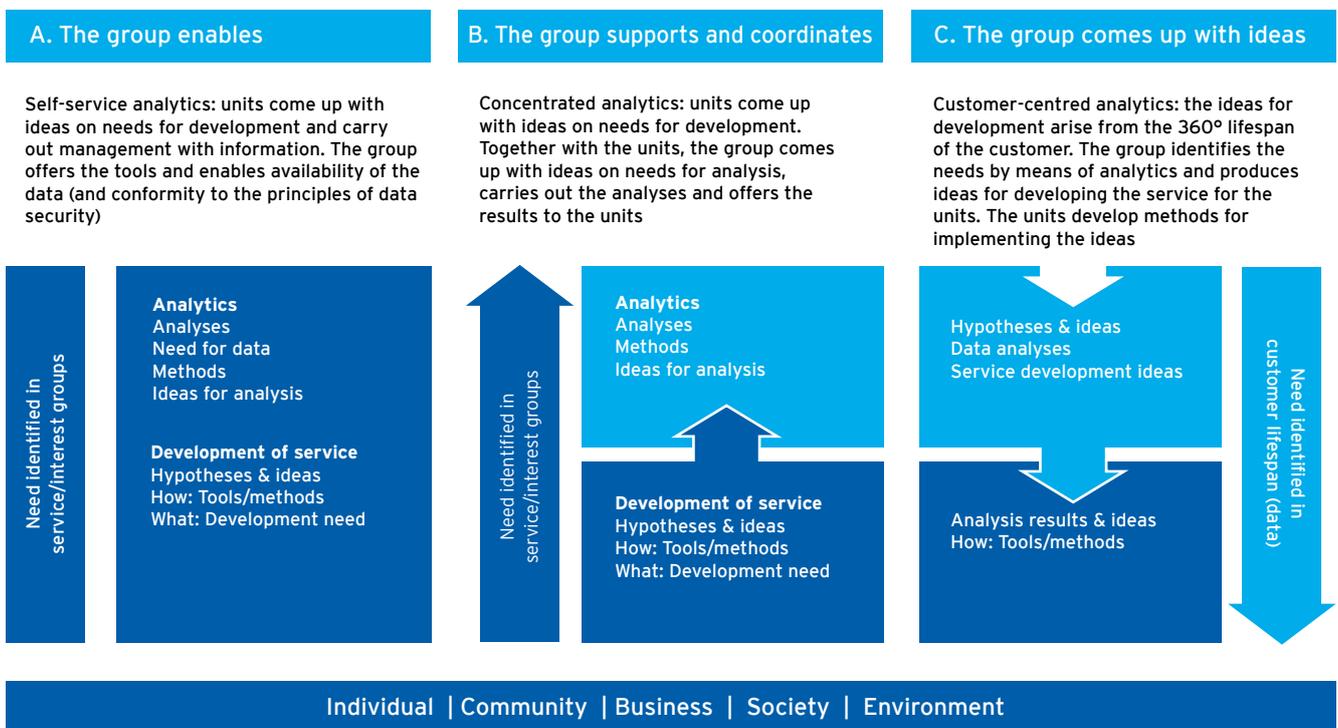
2.1 The operational model for refining data

Refining data and management by data require especially competences related to skills, resources and technologies. These competences can be steered or administered in a centralised way, or they can be decentralised throughout the different organisational units and branches.

A decentralised model enables the refiners of the data to have a good understanding of both the data required for analyses and the needs of business. However, this requires more personnel resources.

A centralised model works well in situations where personnel resources are scarcer. A centralised model also enables an analytic review that crosses the boundaries of organisational units and branches in its search of an overall understanding of customer relationships, and also a more wide-scale ideation based on data.

In practice, a municipality may have more than one operational model in use simultaneously.



The operational models for management by data

The following sections describe the special features of each model from the viewpoint of the city's group management.

The decentralised operational model for refining data

The decentralised operational model means a self-service model for refining data. The various units of the city are responsible for refining their development needs into analytic hypotheses and for carrying out the analyses and the measures based on them. In this model, group management enables the refining of data by offering technical tools for analysis and the steering processes that ensure that the data is processed according to the appropriate principles of privacy and data protection.

The decentralised model requires the branches to have personnel resources and skills for refining data.

The centralised operational model for refining data

In the centralised operational model, the refining of data is coordinated by group management. In this model, group management helps the organisational units or branches to form analytic hypotheses out of business ideas or development needs. After shared ideation, the centralised unit carries out the analyses and produces their results to the commissioning party (e.g. a branch of the city).

In this model, the role of group management is to offer centralised analytic expertise to the branches and units and to provide technical tools for analysing the data. Group management also coordinates the steering processes that ensure that the data is processed according to the appropriate principles of privacy and data protection.

The centralised operational model requires less staff resources than the decentralised model. However, the model requires good cooperation between branches and the centralised unit that refines the data.

The analytic process is interactive in nature. As the process proceeds, analyses are reviewed and findings discussed between the commissioning body and the maker of the analysis.

The centralised data-centric operational model for refining data

In the centralised data-centric operational model, refining the data is centralised, but the developmental ideas are born out of the data, not necessarily out of any business needs identified beforehand. In data-centric development, significant causal connections are sought in the data without preconceptions.

In this model, the group identifies customer needs by means of analytics directly from the data and produces developmental ideas for units and branches. The units develop methods for implementing the ideas.

The strength of this model is highlighted especially when large-scale and often complicated wholes are being examined or when it is wished to predict future patterns of behaviour.

The model is well suited to situations where very large masses of data are processed with cutting-edge analytics (e.g. machine learning and other methods related to artificial intelligence). The model requires deep expertise and wide-ranging usage rights for processing data from several different sources.

2.2 The process of refining data

The process of refining data means the process and measures needed to find a solution to a business need with the help of analytics.

The process has six steps:

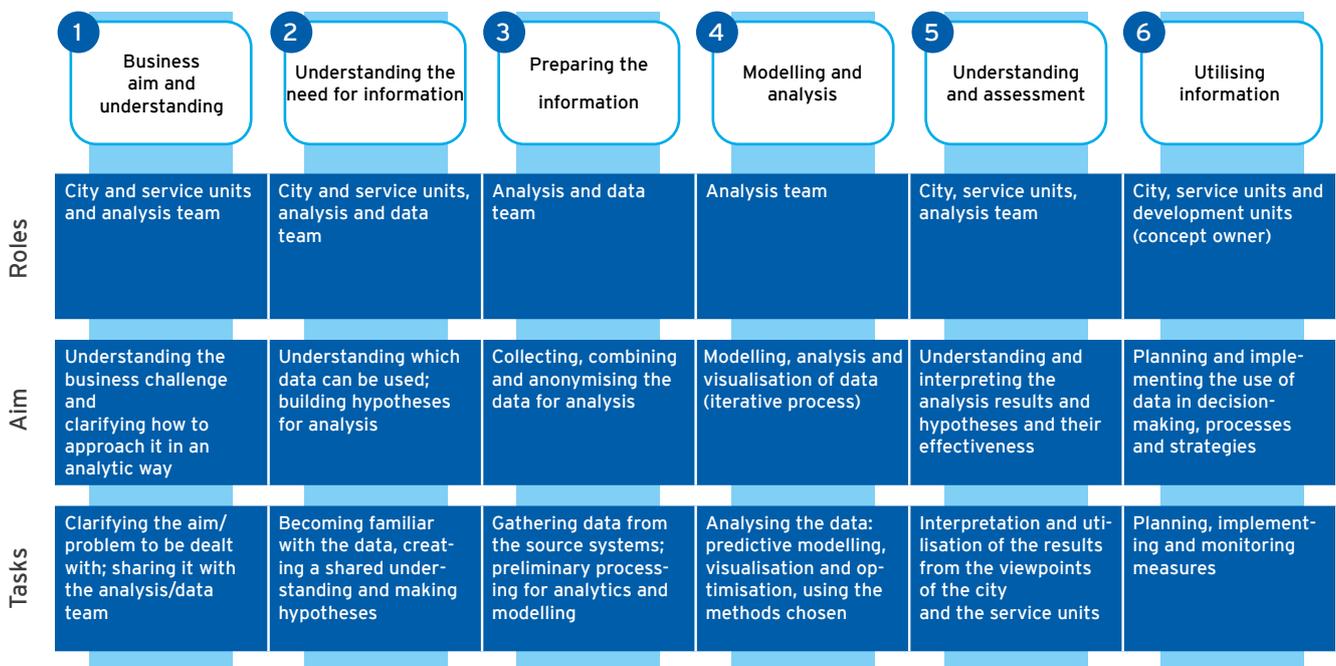
1. Understanding the business need or aim
2. Refining the aim into an analytic hypothesis and a need for data
3. Retrieving the required data and preparing it for analysis
4. Modelling and analysing the data
5. Visualising, understanding and assessing the results
6. Utilising the refined data in decision-making

Identifying the business need or aim is the starting point for the process of refining data. From the aims and needs, questions are deduced to which answers are sought using the data and the refining.

The business questions must be clearly defined at an early stage in the process, but the analysis itself - from understanding the need for information and defining the analytic hypotheses to modelling and visualising the data - is an experimental, iterative process whose phases cannot be exactly defined beforehand.

The final part of the process, utilising the data in decision-making and carrying out the decisions, is management by data. It is thus the phase of refining data that creates value. The analysed information based on data must be combined with other available information. Together the pieces of information create a basis for better decision-making and better measures.

- During the project, the process described in the handbook has been tested in a practical pilot project. This used analysis of data from multiple registers in order to identify the customer paths of the customers of child protective services in Espoo and understand how the native language of the customers affects the customer path.



Process of utilising data in service development

Understanding the business need or aim

The process of refining data always starts from a need. The need may come from within the municipal organisation or from an interest group, but it must be related to business. Without a clear understanding of what the refined data is to be used for, it is difficult to define what should be analysed.

The aim of the first phase of refining data is to identify and understand the business challenge, opportunity or need that can be promoted by understanding data.

This understanding requires cooperation between the party that the need concerns (e.g. a branch of the city) and the maker of the analysis. At this stage, the maker of the must understand what is wanted and to create a preliminary picture of how to approach the matter analytically.

In the pilot project on child protective services, the question was:

“The need for child protective services has increased in recent years, especially for more extensive services. How could the child protective services be developed to better serve both the indigenous population and immigrant families?”

Refining the aim into an analytic hypothesis and a need for data

The next phase of the process of refining data is the formation of analytic hypotheses and questions from the business need. The aim is to understand how the business need can be helped with data and what kind of data this requires.

In the pilot project on child protective services, the preliminary analytic hypotheses were the following:

- How do preventive services and services lighter than placement reach the children of immigrant families (= children whose native language is foreign)?
- How well do these services achieve results for the children, compared to the families whose native language is Finnish or Swedish?
- How much are the different services being used? How does the use of services in foreign languages differ from the use of services in Finnish and Swedish?
- What are the customer paths to the different services? Do the service paths differ by language?

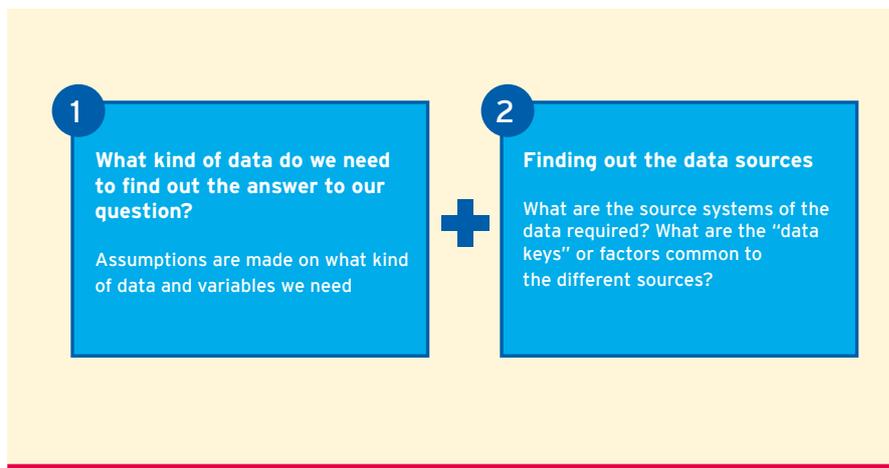
The analytic hypotheses help to answer questions about the data and variables needed. When the needs for data have been preliminarily defined, the data sources required can be identified better.

From the viewpoint of the development of child protective services, it was estimated that data was needed on:

- demographic factors, e.g. gender, age, country of birth
- native language of the children
- native language of the parents
- decisions by child protective services and the services derived from these

The time frame of the research materials was defined as seven years (2010-16). It was estimated that the data would give a comprehensive picture of the customer paths and service volumes in child protective services.

On the basis of the data descriptions, it was found that the required data is available by combining data materials from the SosiaalEffica customer information system and the population information system. The data were combined based on the personal number.



Understanding the need for data

Retrieving the required data and preparing it for analysis

Retrieving and preparing the data is intended to generate the data required for the analytic task. At the same time, the preconditions set by privacy and data protection must also be taken into account.

If the analysed data includes personal data, the EU's General Data Protection Regulation requires that the data be anonymised as well as possible to prevent natural persons or information on them from being identified directly or indirectly.

In this phase, the data is retrieved from the source systems, and the combinability of the various pieces of data is checked. If required, the data is anonymised and pseudonymised, and the data is then saved to the analytic environment (e.g. a data reservoir). As the analysis of data is experimental and iterative in nature, the retrieval and preparation of data must often be supplemented as the process progresses.

The data analyst may participate in defining the data required, but for reasons of data security, the analyst must not participate in retrieving, combining and anonymising the data.

In the pilot project on child protective services, the data experts of the City of Espoo retrieved the required data materials from the two source systems. They combined the pieces of data on the basis of the personal number and then anonymised the data material with the SAS Enterprise Guide tool. The anonymised data was handed to the analyst for further processing.

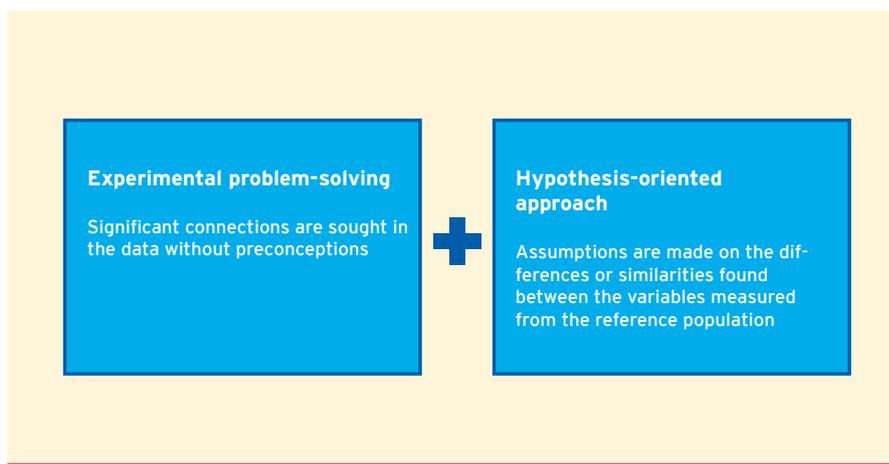
Modelling and analysing data

The analysis of data aims to find answers to business needs in the data materials. In the analysis, data from different sources is modelled with statistical, mathematical and other methods, creating a better understanding and new findings or viewpoints.

The analytic approach is selected according to the research aims. The approach can be hypothesis-driven, in which case the analyses are based on verifying predefined hypotheses, or it can be experimental. In an experimental approach, significant connections are sought in the data without preconceptions. Mutual dependencies can be identified that were unknown before the analysis started.

The approaches are not mutually exclusive, but complement each other instead. Modelling a particular object of analysis may well make use of both approaches. The methods of modelling data are described in more detail under “Methods of analysis and modelling” in section 2.3.

When the data is modelled, new deficiencies and errors are also often found in it. These findings often lead to improved data quality in the long run.



Modelling and analysis: approaches

The starting point for the pilot project on child protective services was a hypothesis-driven approach that tested the following:

- how do preventive services and services lighter than placement reach the children of immigrant families (= children whose native language is foreign)
- how well do these services achieve results for the children, compared to families whose native language is Finnish or Swedish
- how much are the different services being used, and how does the use of services in foreign languages differ from the use of services in Finnish and Swedish
- what are the customer paths to the different services, and do the service paths differ by language

In connection with the modelling, deficiencies and qualitative challenges were detected in the data and its different variables. These led to new data requests, new data classifications and the creation of new variables. The definitions in the basic data (the so-called master data) were not the same in both systems, which made the processing and analysis of the data more difficult.

In part, the findings detected in the data also led to improved recording practices. This will improve the quality of data in the long term.

Visualising, understanding and assessing the results

The aim of visualising and reporting the results of analysis is to put the refined data in an easily understandable and user-friendly form. The owner of the business need (the utiliser of the data) must be able to review and interpret the findings; to assess the results of analyses and the functionality of analytic hypotheses; and to draw conclusions from the findings to support decision-making.

Visualisation of the refined data can be done with traditional tools such as PowerPoint or Excel. However, more modern tools for visualisation and reporting, such as the Microsoft PowerBI used in the pilot project on child protective services, often enable a more user-friendly and versatile view of the results of analysis. These tools enable not only drilling deeper into the data and findings but also new viewpoints and visualisations.

Utilising the refined data in decision-making

Management by data requires that the final phase of the process of refining data - utilising the data - leads to concrete measures being taken. The benefit from the previous phases of the process will be rather limited if the findings received from the data are not utilised in decision-making.

The decisions may be small ones related to simple daily routines, or large ones involving changes of the entire strategy or focus.

Refined data has many uses, such as the following:

- support for decision-making
- predicting and simulating the future
- monitoring and assessment
- changing operational processes
- increasing attractiveness
- creating new preventative services
- focused communications
- innovating
- measuring performance

2.3 Cases of the use of data; users of the data

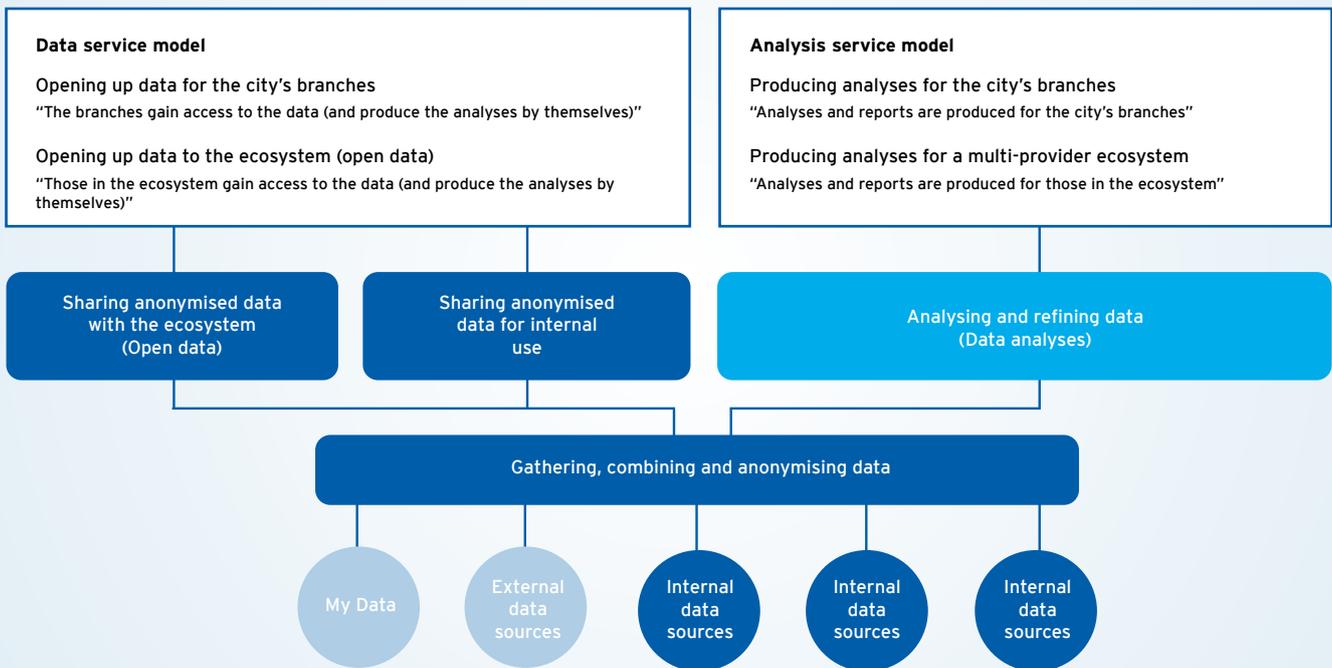
Cases of the use of refined data can be divided in two categories:

- cases based on refined data, and
- cases based on the end results of analysed data.

Cases based on refined data

In these cases, there is a need to produce tailored data material (e.g. one compiled from several data sources and anonymised) that the party needing information can refine further by means of analytics. For instance, this kind of need is found among interest groups, such as researchers who want to combine the received data with data from other sources and create research hypotheses of their own. The need for this kind of use can also come from within the municipal organisation, from units that are themselves able to analyse the data and to refine it further.

Cases of the use of refined data are partly similar to cases of the use of so-called open data available to all, but differ in that the data materials are tailored for a particular purpose. The need for tailored data is also often one-off in nature, which is not the case with open data.



The general process of refining data and cases of use

Cases based on the end results of analysed data

In cases based on the end results of analysed data, the end user of the data needs a ready-made analysis: a report or visualisation prepared on the basis of research hypotheses. The user of the data does not require expertise in refining data, instead having direct access to the findings produced by the analysis.

Cases based on this model come from within the municipal organisation and from those external interest groups that have no expertise of their own in refining data. Among other things, these include non-profit associations and NGOs.

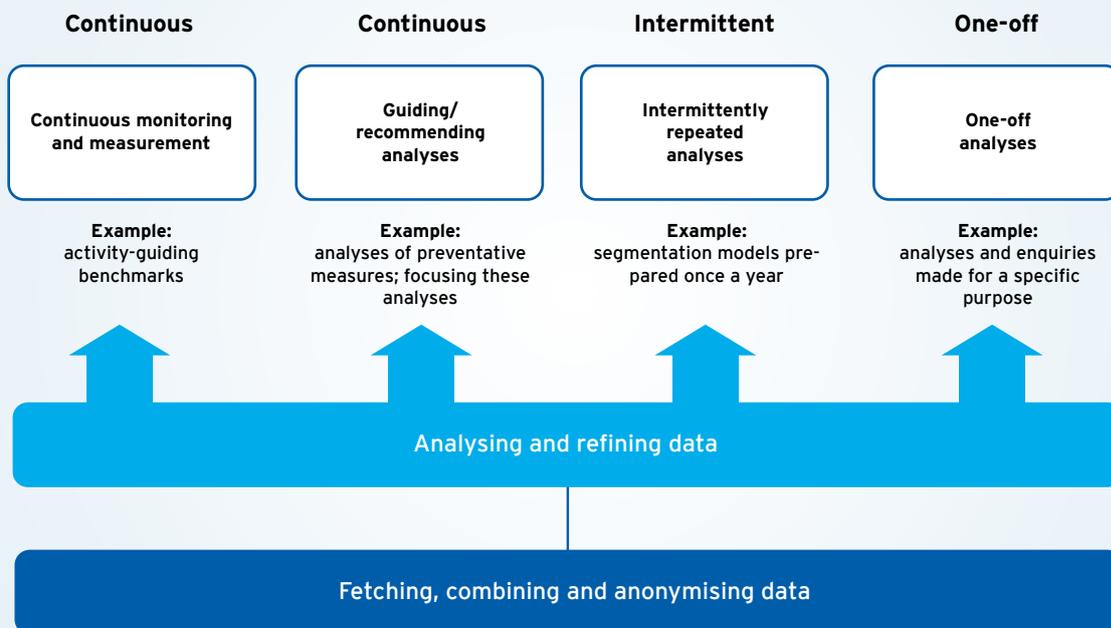
Cases based on analyses are need-driven and, therefore, take many forms, but they can be divided in

- continuous,
- intermittent or
- one-off.

Among other things, continuous analytic services include indicators for continuous monitoring, such as yardsticks for the different aims of an activity, or cutting-edge analyses intended for operative use that produce recommendations or suggestions. The latter category includes analytics related to the automation of marketing and analyses having to do with machine learning.

Some of the analyses are carried out more intermittently, such as customer segmentations. Some are done only once to meet some special need.

The extent to which the need for analysis will be repeated is a central consideration when the implementation and repeatability of an analysis is being designed.



The general process of analysed data and cases of use

Users of data

The users of data are the groups of people the refined data is offered to. The users of data often come from the group or other party that has identified the need that the refining was based on. In some cases, a larger group of persons can also belong to the users.

The users of data can be grouped in:

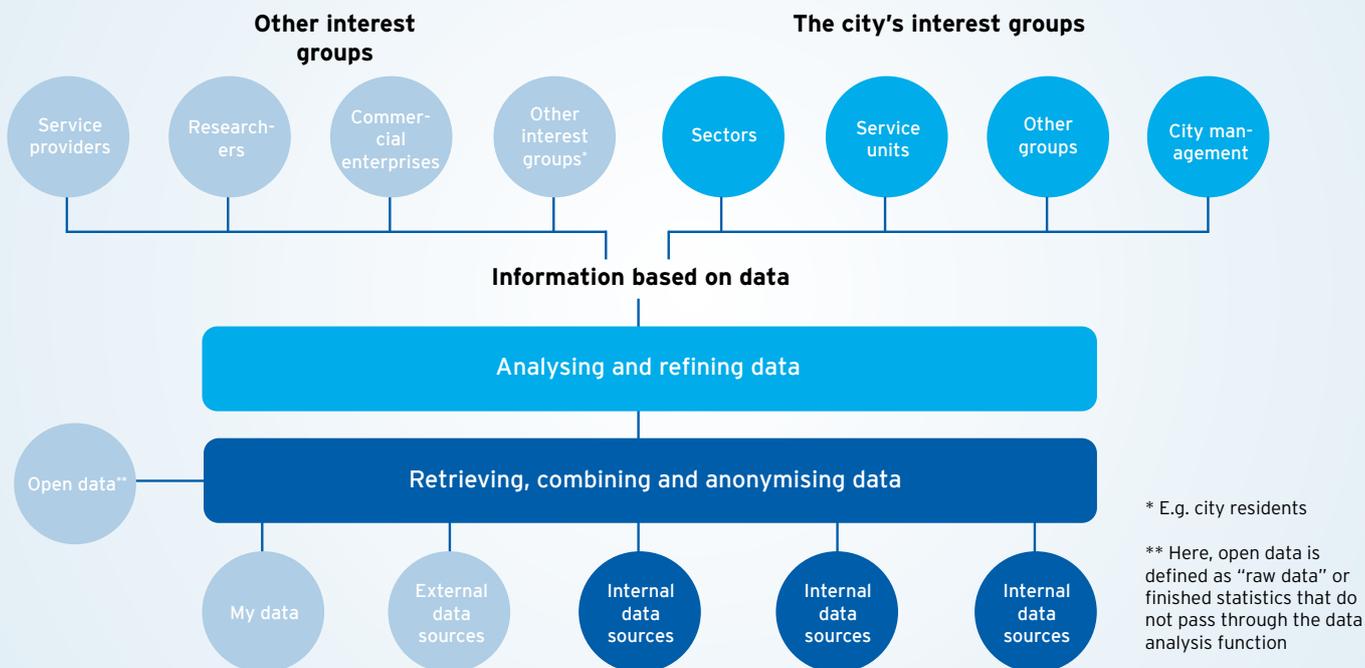
- users within the municipal organisation, and
- users outside the municipal organisation.

Users within the municipal organisation include

- branches
- service units and
- the city management.

Users outside the municipal organisation include representatives of different interest groups, such as

- service providers
- universities and researchers
- non-profit associations and NGOs, and
- commercial actors.



Users of data

The information content of the analyses must be produced in the form that is most suitable for the end users. The same data can thus be produced in different forms for different users.

Means of publishing the data by role

When analyses are published for end users, both the purpose of use of the analysis and the user profile must be taken into account. The same data can be disseminated refined in different ways to different user groups, depending on their information needs, skills and time pressures. Designing the right means of publication is an important part of the process of refining data, as it affects the understanding of the data and, therefore, also the usability of the data.

Information can be published in the following forms among others:

Standard reports

The content and layout of the reports are ready-made. The content of the report changes dynamically, e.g. by date, removing the need for users to select delimitation by data when the report is run. Example: activity-guiding benchmarks

Example of user: City management

Dynamic reports

Like standard reports, but users can drill into the data and create predefined delimitation conditions for the information to be reported, e.g. by time or by cost centre number.

Example of user: Branch management

Alert reports

Reports on continuous monitoring that follows the deviations and changes. Automatic alerts help by drawing attention to deviations.

Example of user: Service unit

Ad hoc reports

Reports and analyses done in response to a one-off need. Often dynamic reports that enable drilling into the data, but can also be implemented with traditional tools such as PowerPoint or Excel.

Example of user: Service unit

Self-service BI

Reports that enable focusing all the way down to the raw data. Enable users to create their own reports, to search for causes and effects and to examine the data in a versatile way. Require expertise in both tools and data.

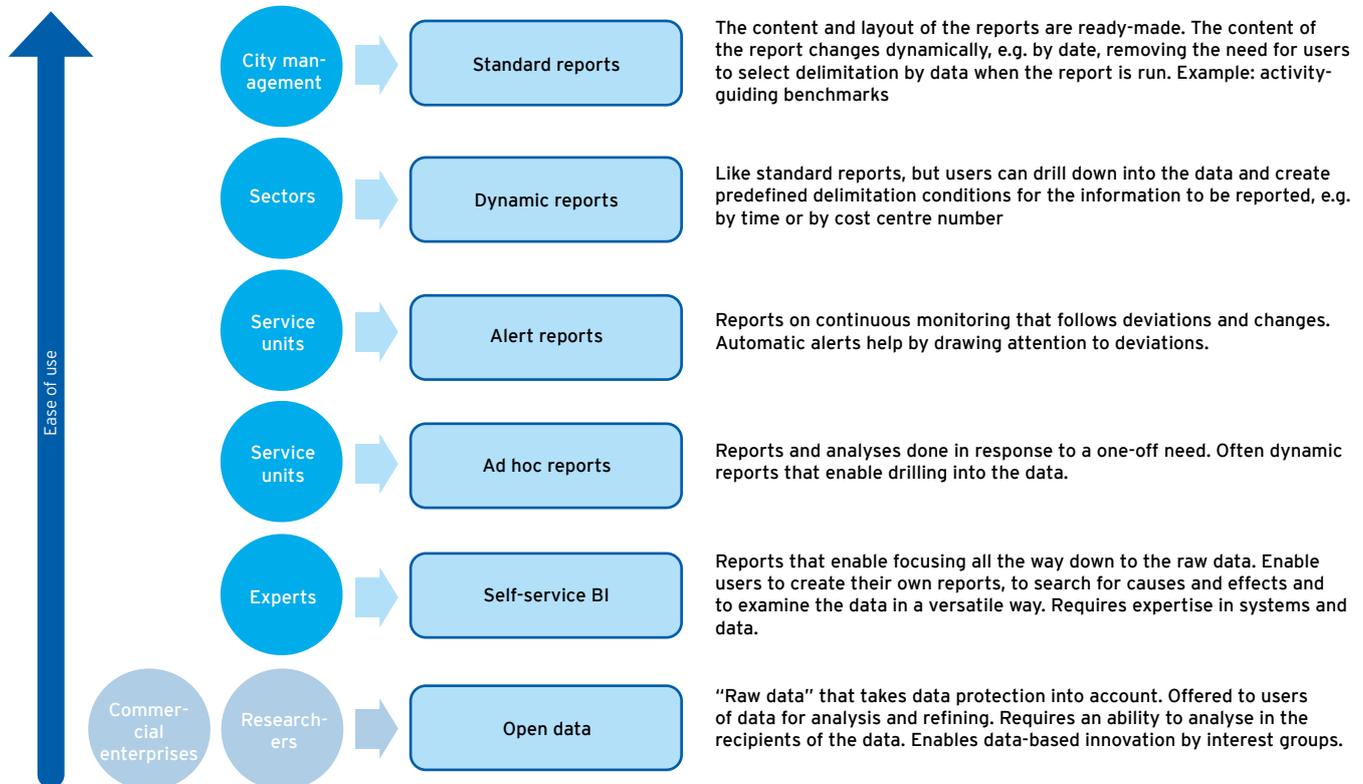
Example of user: Service unit expert

Tailored or open data

"Raw data" distributed in the form of files, taking into account the GDPR. Offered to users of data for analysis and refining. Requires the recipients of the data to have an ability to analyse. Enables data-based innovation by interest groups, for example.

Example of user: Researchers, commercial actors

In addition to the cases of use mentioned above, refined information can also be disseminated directly into other data systems. It is often a case of refined data that is needed to steer operative processes or to make them more efficient. Examples are the automation of marketing or the responses and recommendations given by chatbots on the basis of data.



Means of publishing the data by user role

Methods of analysis and modelling

The previous section described how information can be distributed to end users. However, the method of publication is independent of how the data has been refined or analysed before being published.

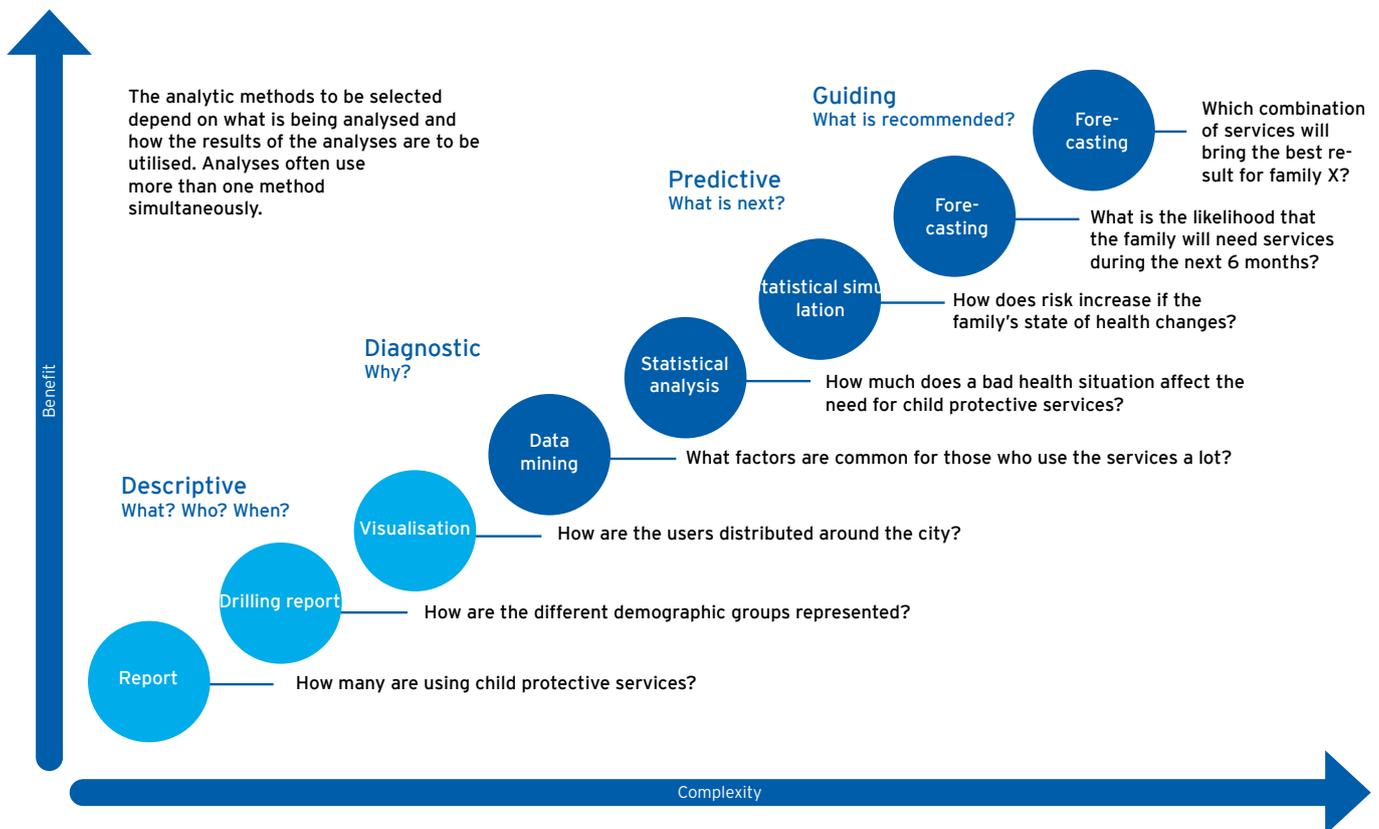
There are many different methods of analysis and modelling, but they can be roughly divided in two categories:

- Business Intelligence, meaning traditional reporting and visualisation that intends to describe what has happened, and
- analytics, which is often forward-looking and aims to understand why something happens and to predict, simulate and steer future events.

The analytic methods to be selected depend on what is being analysed and how the results of the analyses are to be utilised. Analyses often use more than one method simultaneously, and several different approaches can be used in solving the same problem.

From the viewpoint of the management of competences, it is important that the analyst can identify different ways of approaching a research hypothesis. In this way, the analyst can change the modelling method on the basis of the findings made, if this is required.

Different methods of analysis also require different types of expertise. The more developed the analytics to be used is (e.g. machine learning or artificial intelligence) the deeper the expertise in mathematics and statistics that will be required will be.



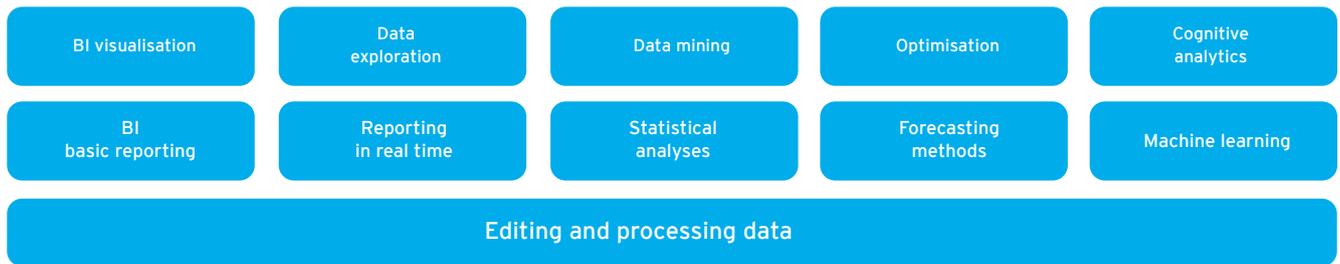
Defining the data needs and data sources for refining data

Refining data can thus be seen as a model with three levels: utilising information, producing information (analysis) and managing information. The choice of components is guided by business needs and the resulting need to make use of refined data.

Utilising the data (method of publication of analysis results)



Refining the data (method of analysing the data)



Data management



The levels of utilising and refining data

2.4 Resources and competences

The management, production and dissemination of data and the management and development of the entire process of refining data often require new kinds of roles and responsibilities in the organisation.

Roles and competence profiles

One person may have several competence profiles, but the profiles required the most often are the following:

- sponsor or leader of analytics
- expert
- data analyst or data scientist
- visualisation expert, and
- more technical expert or data architect.

The following figure describes the most typical roles and their related competence profiles.



Sponsor/leader of analytics

- Strong understanding of business combined to experience in applying analytics to solving business problems. Experience in leading teams of experts.



Data and analytics expert

- Understanding of business combined to practical experience in applying analytics to solving business problems.
- Practical understanding of analytic methods, data processing and information systems



Data analyst/Data scientist

- Good skills in acquiring and modifying data, modelling, statistics, machine learning, developed analytics or mathematics.
- Capable of solving problems in modelling of algorithms, testing, processing, etc.



Visualisation expert

- Experience in visualisation of information, graphic presentation of results, researching information and drilling into data
- Strong experience in visualisation tools (e.g. PowerBI, QlikView, D3.js, etc.)



Data architect/Technical expert

- Experience of data ETL (extract, transform, load) from different databases, e.g. SQL, NoSQL, HDFS
- Experience of examining the quality of information, the anonymisation of data and different data management tasks

The roles and competences needed in refining data

Areas of responsibility

Different responsibilities are also related to different task descriptions. The areas of responsibility must be defined according to the operational model of each organisation, but generally speaking, a few critical roles can be highlighted.

Sponsor or leader of analytics

As management by data may entail changing the entire culture of management of an organisation, it is important for the activity to be supported by the management and that a representative of the management is participating in managing the change.

The sponsor or leader of analytics is a kind of “interpreter” who helps different actors in the organisation to understand the opportunities and preconditions of management by data. The sponsor also helps to implement the changes required in ways of acting.

Analytics expert or analyst

The analytics expert fits together business needs and analytic methods. The responsibilities of the analytics expert include understanding business needs and facilitating communication between business and analytics. Often the expert is also responsible for conducting analyses and interpreting the results together with business.

Data scientist

There is also an increasing need for experts in developed analytics, such as analytic innovations, machine learning and modelling of algorithms in artificial intelligence. The responsibilities of the data scientist also include testing new analytic methods and algorithms as part of analytic problem-solving. The data scientist is responsible for the selection and functioning of analytic models.

Data architect/technical expert

Data architects and experts in technical data processing also have an important role. Administratively, these persons often belong to the data management organisation, but from the viewpoint of refining data, their role is central. Their responsibilities include preparing data for analysis: developing and constructing the retrieval, merging and anonymisation of data and the related automation.

The responsibilities also often include the development of data reservoirs and the construction of database connections to different source data systems.

Sponsor/leader of analytics

- Challenging traditional models and innovating the new. Bringing into view the opportunities of analytics in the development of customer management and services. Responsibility for the steering and supervision of analytic activity.

Data and analytics expert

- Fitting together business needs and analytic methods. Facilitation of communication between business and analysts. Responsibility for making analyses and interpreting the results together with business.

Data analyst/Data scientist

- Analytic innovation. Testing and applying new analytic methods and algorithms as part of analytic problem-solving. Innovating and developing analytics and analytic models. Responsibility for the selection and functioning of analytic models.

Visualisation expert

- Development of the reports and visualisations related to data and the results of analysis to meet the needs of end users. Innovating new reporting models and developing the usability of reports. Responsibility for the visual usability of analyses.

Data architect/Technical expert

- Developing and constructing automation of data retrieval, merging and anonymisation. Developing the reservoir of data and building database connections to different source data systems. Responsibility for the availability of data.

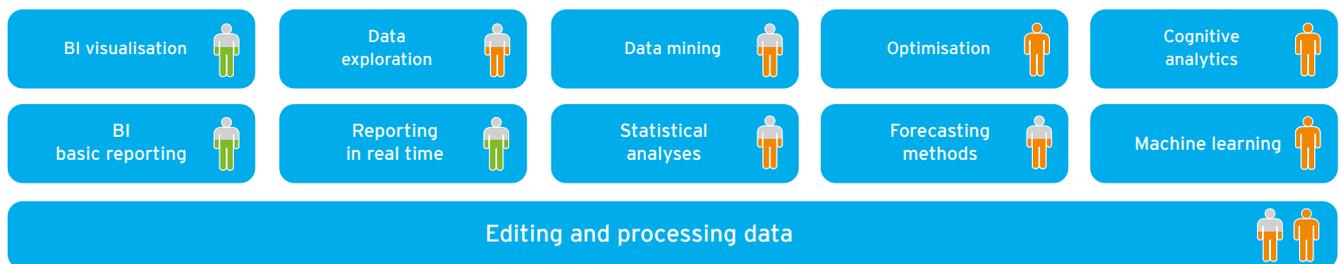
The roles needed in refining data and their areas of responsibility

The roles and areas of responsibility are reflected in the three-level model of refining data as follows.

Utilising data



Refining data



Data management

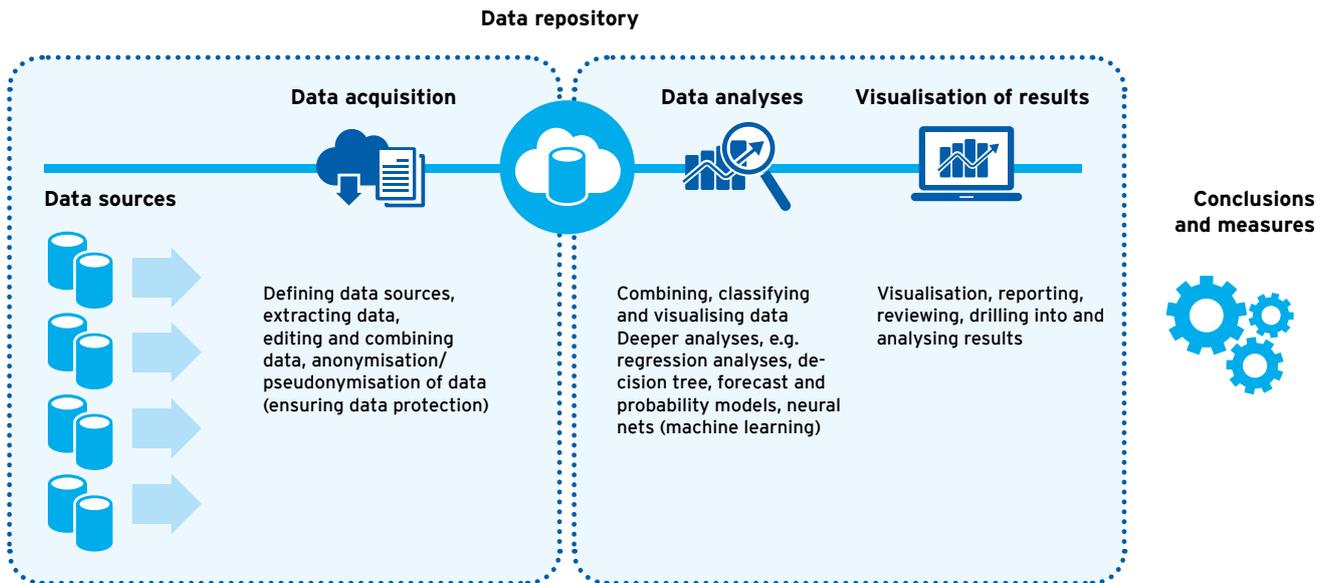


The roles needed in refining data

2.5 The technology of refining data

As an upper-level technical description, refining data requires technologies for the retrieval, storage and analysis of data and the visualisation of results.

The technologies of data retrieval and storage are often tools of data management, while the technologies related to analysis and visualisation are within the area of responsibility of refining data. Technological solutions often contain functionalities that cover many different areas, but as a whole, refining data requires more than one tool almost without exception.



Technical description of the upper level of refining data

When the tools and technologies are being selected, both the needs of information technology and the needs of business must be taken into account. Technologies related to analysing data are developing quickly. New alternatives come on the market almost monthly, and different tools have a lot of overlap in their functionalities. For this reasons, the selection of technologies should be approached with appropriate seriousness.

A good way of proceeding is to form an understanding of the analytic needs identified; to chart the potential new needs as well as possible; and to form a gradual model of proceeding on the basis of these. A gradual model will offer good guidelines on the technologies needed at different stages in the path of development.

The following image shows a few of the more central alternatives among the technologies intended for the different phases of refining data. The list is not exhaustive.

2.6 Data protection and privacy in data refining

Refining and analysing data always requires the processing of data materials. The data can be structures or unstructured, numeric or textual, or the analysis may

be based on the recognition of images or speech. When the aim is to analyse data that can be used to develop customer management and services, it is natural that part of the data to be processed is so-called personal data.

The EU's General Data Protection Regulation (GDPR), which came into force on 25 May 2018, defines personal data as follows:

“‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person”.

According to the GDPR, then, personal data means all the information from which a natural person (a human being) can be identified directly or indirectly. Among other things, this information includes a person's name, telephone number, personal number and photograph.

If the analysed data contains personal data, it must be anonymised as part of the preparation as well as possible to prevent natural persons and their information from being identified directly or indirectly.

Anonymising the data

In anonymised data, all the information enabling persons to be identified is redacted and cannot be restored. For instance, in anonymised data, the personal number is replaced with an individual identifier, enabling individuals to be identified on the different rows in the data, but not by name. The required categorical variables (gender, age group, etc.) can be combined with the identifier if use of the data requires this. Anonymous information serves well in cases where there is a need to understand the behaviour and wishes of customers without identifying them as persons.

Pseudonymising the data

Pseudonymising the data is another way of hiding personal data from the analysed materials. In pseudonymisation, the identifying information in the combined data is replaced with material-specific information on individuals that is saved in such a way that only particular persons (e.g. those carrying out anonymisation) have access to it. The pseudonymisation of data can be considered in cases where there is likely to be a need to connect the end result of an analysis to a person, e.g. when preventive services are to be focused on persons belonging to an at-risk group. However, it must be taken into account that pseudonymised data has the same status as unprocessed data from the viewpoint of the GDPR.

Data protection and processing the data for analysis

The essential thing is that the acquisition, entry and merging of information is done by persons who do not make use of the data in their own work. For this purpose, a dedicated and separate technical or logical environment (data reservoir) is defined, and the combining, anonymisation or pseudonymisation is carried out there.



The division of labour in refining data with data protection as a starting point

The persons analysing the data must not have usage rights to any unanonymised data or any technical opportunity to process or see pure personal data.

By dividing the environment in two separate wholes in terms of usage rights (a sensitive and a non-sensitive one), it can be ensured that the utilisers of the information (those carrying out analyses) do not have access to any data that is sensitive in terms of data protection. In this way, the protection of sensitive data from the utilisers of the data can also be ensured.

Ready-made rule sets and automation can be used in combining or anonymising data if this is required, e.g. in continuous analyses.

Some analyses, such as steering or recommending analyses, may require the identification of individuals. For instance, focusing preventative measures on particular individuals requires them to be identified. It is useful to identify these analyses in an early stage of the process of refining data. This will leave time for preparations and enable appropriate processing.

	Continuous monitoring and measurement	Guiding/recommending analyses	Seldom repeated analyses	One-off analyses
	Example: activity-guiding benchmarks	Example: analyses of preventive measures and of focusing these	Example: analyses done once a year	Example: analyses and reports made in order to meet a particular need
Analysis and reporting of data	High degree of automation in the implementation of analyses, reports and visualisation	Integrating analyses in service production, e.g. in the form of recommendations	Saving the models of analysis and reporting and enabling them to be reused	Topic-specific analyses and reporting models produced as needed
Anonymisation and storage of data	Automation of anonymisation. Connection back to persons cut off. Continuous storage of data	Pseudonymisation of data; identifying information replaced with material-specific information on individuals, to which only designated persons have access	Anonymisation or pseudonymisation of data depending on need. Storage of data as needed	Anonymisation or pseudonymisation of data depending on need. Storage of data as needed
Retrieving and combining data	Automated process	Primarily an automated process	Primarily a manual process	Manual process (making use of tools)
	High degree of automation of anonymisation	Connection back to persons must be possible	Saving and reusability of models	Low degree of automation in the production process

Data protection and the ways of using the data

The permissions process for the refined data

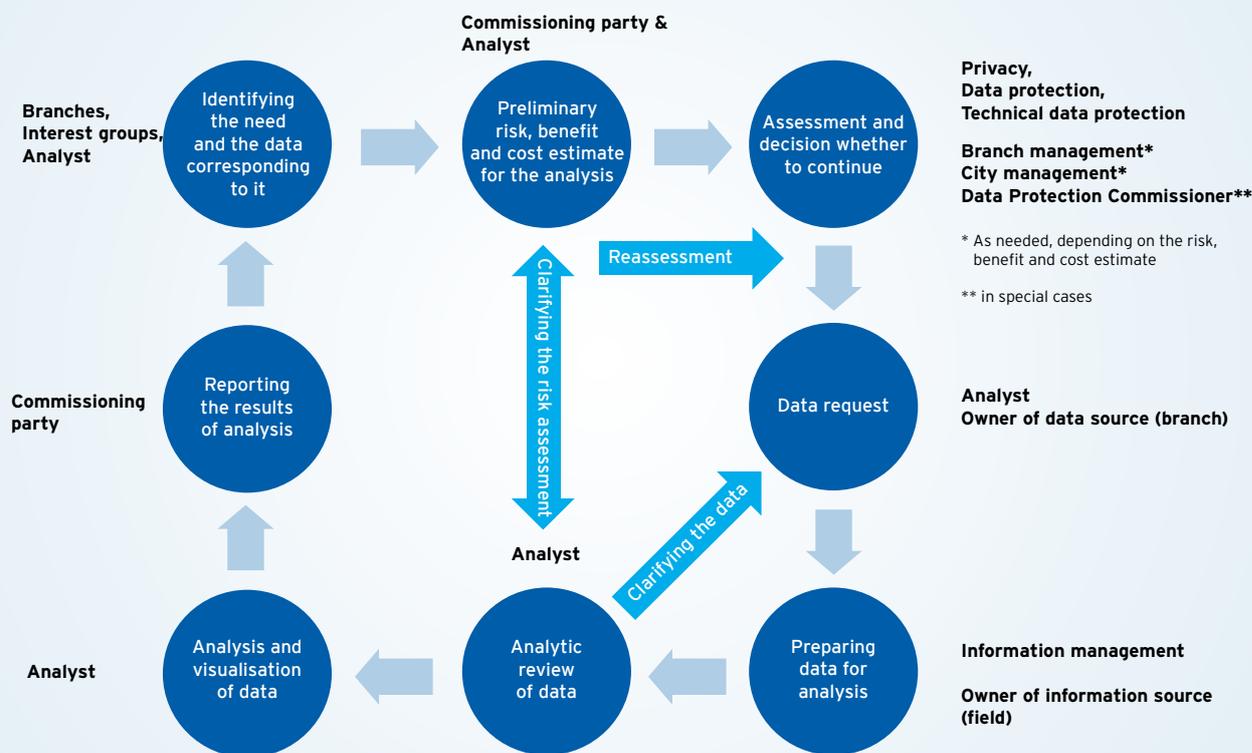
In spite of anonymisation, there is always a risk of natural persons or their data being identified. For instance, in the future, this risk can be realised by combining different pieces of data and the information they contain and thus creating new means of identifying the information in the original data. Future technological developments (e.g. artificial intelligence) may also enable the use of more efficient statistical algorithms, increasing the risk of identification.

In addition, business needs may call for the analysis of unanonymised data.

To ensure the processing of analyses according to the requirements of privacy and data security, a description and set of guidelines on the measures and inspections carried out during the process of refining data has been created in cooperation with the City of Espoo. These are used to prepare an assessment of the benefits, risks and possible costs of each analysis, which is used in deciding whether an analysis can be carried out.

The process has eight main phases:

1. Identifying the need and the data corresponding to it
2. Preliminary risk, benefit and cost estimate for the analysis
3. Assessment and decision whether to continue
4. Data request
5. Preparing the data for analysis
6. Analytic review of the data
7. Analysis and visualisation of the data
8. Reporting and utilising the analyses



Process description, phases of permissions process

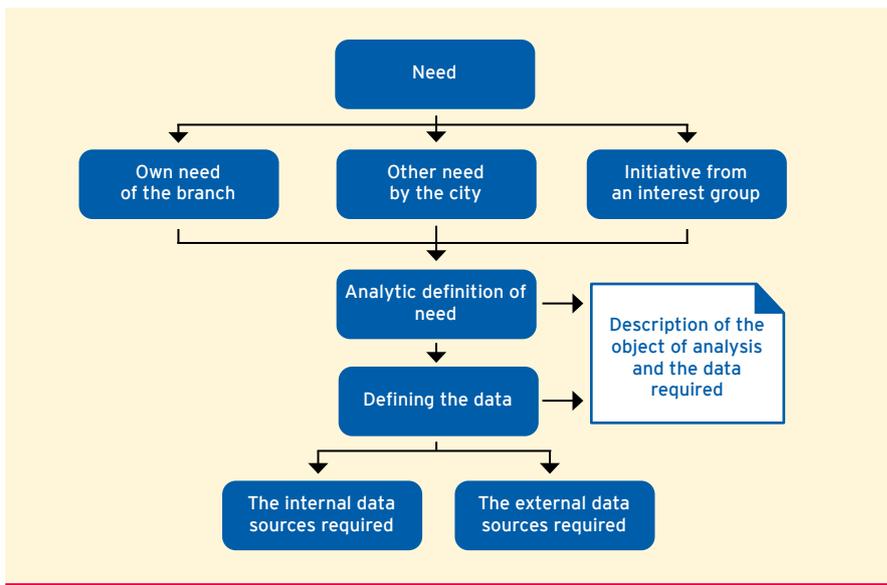
Identifying the need and the data corresponding to it

The impulse for analysing data can come from very many different sources, but as a rule, the analysis is always prepared for some purpose. Many branches already function actively in different cities and try to find objects of analysis on their own initiative. Nevertheless, the municipal organisation should also include a centralised functionality that promotes analytics and management by data and proactively supports initiatives across branch boundaries.

The initiative and request for a data analysis can also come from outside the municipal organisation (a business, research institution or non-profit association) or from a branch that needs to combine data from several different branches.

Regardless of the source of the initiative, the need must be defined in an analytical hypothesis that defines and describes (at the upper level) what must be understood or known to make the analysis correspond to the business need involved. The first estimate of the data required for the analyses is derived from this analytic hypothesis. Defining the data will also give a preliminary estimate of the data sources needed. Data sources can be internal or external. There can be one or more sources.

The aim and content of the analysis and the data and data sources required are described and documented.



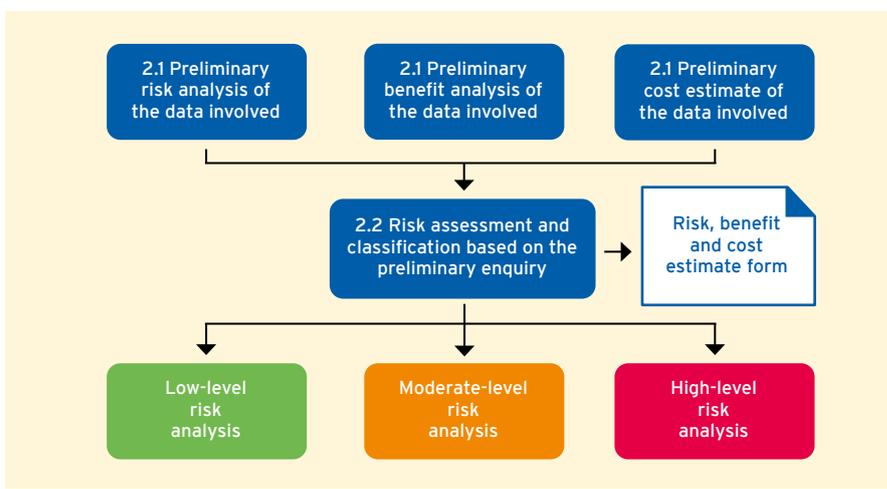
Process description: phase 1 of the permissions process

Preliminary risk, benefit and cost estimate for the analysis

For each initiative, the party making the initiative prepares an assessment together with the analyst, examining the risks and costs related to the analysis and potential combination of the data and the benefits related to utilising the results of the analysis. The risk assessments consider both the vulnerability to risk of the so-called “raw data” and the risks of the data to be analysed and reported. If the analytic need requires the data to be processed to contain unanonymised personal data, the risk of the analysis is always assessed as high or very high.

In the same connection, the costs caused by processing the data (e.g. anonymisation, quality control, IT costs, outside costs) and maintaining it (e.g. updating frequency) is always assessed as well.

The estimate is done by filling in a separate estimate form (Appendix 1). As the data has not yet been examined at this phase, the document is updated as needed (see “Analytic review of the data” below). The completed assessment forms and the preliminary assessment of the risk level of the analysis are primarily submitted for review by the persons responsible for privacy, data security and technical data security.



Process description: phase 2 of the permissions process

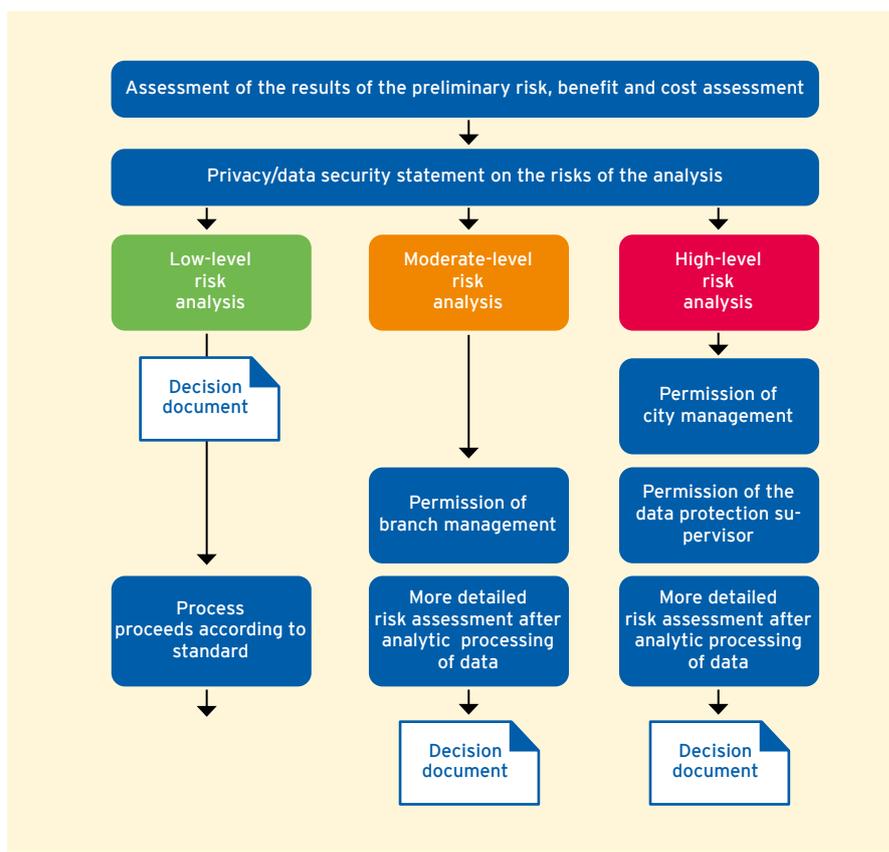
Assessment and decision whether to continue

The party that has carried out the risk, benefit and cost analysis takes the results of the risk analysis to an assessment group that consists of representatives of data security and technical data security.

On the basis of the documentation, this group makes a preliminary assessment of the risks related to the analysis and issues a statement on them.

- If the risk level of the analysis is defined as low, the process will continue according to the standard model.
- If the risk level of the analysis is defined as moderate or cannot be assessed clearly, the analysis is taken for separate handling by the branch management and risk management of the commissioning party. If it is decided to continue the process, work continues normally until the analyst has a clear view of the content of the data and the analyses performed on it. A new risk assessment is then made.
- If the risk level of the analysis is defined as high, permission for the analysis is sought from city management as well as from the Data Protection Commissioner, if required. If it is decided to continue the process, work continues normally until the analyst has a clear view of the content of the data and the analyses performed on it. A new risk assessment is then made.

The decisions made on analyses with a moderate or high level of risk are documented separately. The decision document for a low-risk analysis is the original description and risk assessment form approved by the person responsible for data security.



Process description: phase 3 of the permissions process

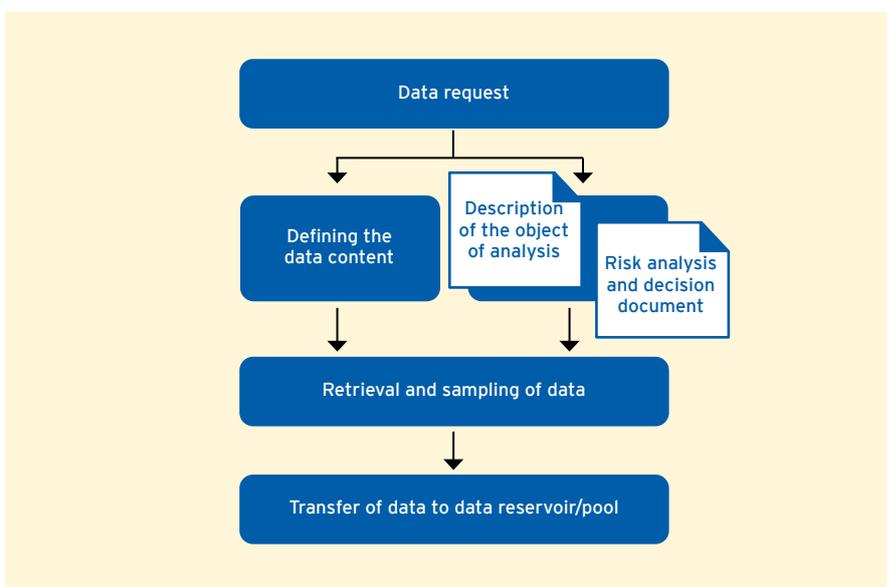
Data request

At the data request phase, the analyst submits the analysis description and the statement of the person responsible for data security/privacy on the risks of the analysis to branch management and to the person responsible for the source system.

The analyst and the system expert define the data and data fields required from the system. If data is needed from several different systems, the analyst repeats the same procedure with all the systems or branches involved.

The systems expert picks the required data and data fields from the source systems. Secure techniques of data transmission is used in the picking and transfer of data. If required, the techniques used are described in the analysis document.

If the data is estimated to contain personal data or any material referring to such data, the analyst does not participate in the retrieval of data or the handling of the "raw data" in any way. After the data is retrieved and picked, it is transferred securely to the data reservoir or pool for anonymisation.



Process description: phase 4 of the permissions process

Preparing the data for analysis

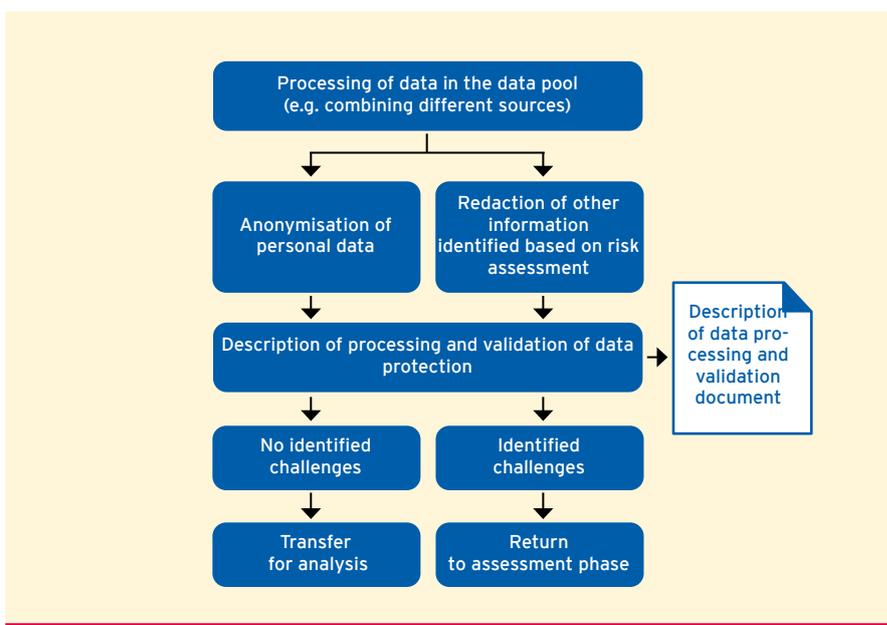
During the phase of preparing the data for analysis, municipal IT experts familiar with data security and anonymisation are responsible for measures related to anonymising the data. They also carry out tasks such as the technical merging of data. If the data contains personal data, its technical processing involves an anonymisation that takes into account both national legislation and the EU's General Data Protection Regulation.

The aim of anonymisation is to produce data that:

- cannot be used to identify individual persons directly or deductively
- cannot be used to connect information on an individual person to the anonymised material
- cannot be used to draw conclusions from an individual person.

The anonymisation must take into account all the means of identification that can reasonably be expected to be used. In addition, identification by both the possessor of the original data and by third parties must be prevented. The process of anonymisation must be irreversible.

Any other information identified on the basis of the risk assessment is also redacted at the same time. The technical processing of the data is described, and the results of the processing are validated. If required, anonymisation measures are continued until the data material conforms to the General Data Protection Regulation.



Process description: phase 5 of the permissions process

Analytic review of the data

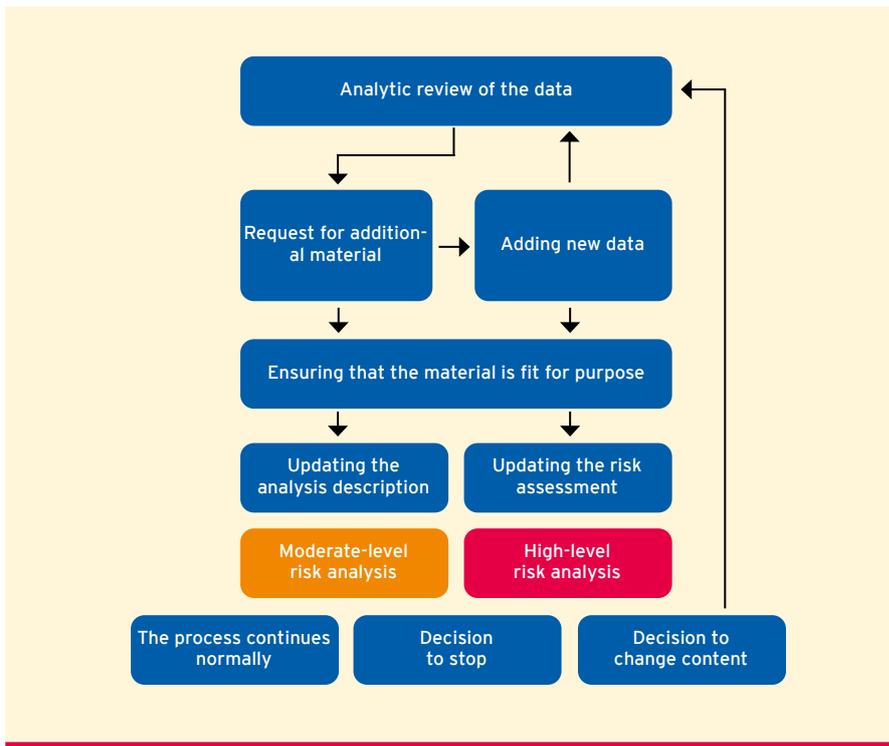
The analytic review of the data ensures that the data content corresponds to need and that the data is suitable for the desired analytic procedures and testing of hypotheses.

The review is often carried out iteratively, in phases. Possible deficiencies or problems are only detected when the first analyses are performed. It is normal to turn back to the data sources and complete the required initial data more than once.

When the data is reviewed from different viewpoints and the data material received is found to meet the needs of analysis, it is time to add to the document describing the analysis and to clarify the risk analysis as required.

If the risk of the data analysis is classified as moderate or high, the completed documents must be taken for reassessment by the inspection group for each initiative. Based on the clarified description, the inspection group will reassess the initiative and make a decision according to which:

- the analysis continues according to the normal process
- the analysis is rejected, or
- the content of the analysis is redefined into a form that is safer from the viewpoint of risk.



Process description: phase 6 of the permissions process

Analysis and visualisation of the data

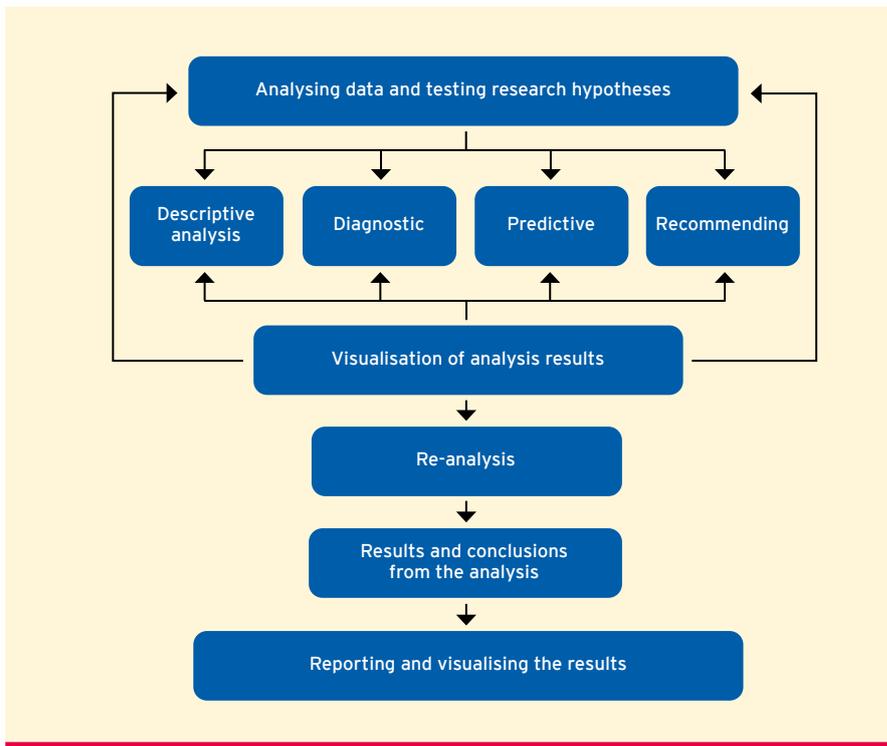
When the analysis and data content are approved, the data analysis proper can begin. The data analysis is performed by an analyst. Persons participating in the technical merging and anonymisation of data do not participate in analysing it.

Familiarisation with the data in the previous phase has already enabled a good understanding of the data content. The tasks in this phase concentrate on testing analytic hypotheses and interpreting the results.

Depending on the research hypothesis, the methods used in analysis can be descriptive, diagnostic, predictive or recommending. Depending on the object of analysis and the research hypothesis, more than one method can also be used simultaneously. Analyses almost always contain a section devoted to descriptive analytics. The results are also observed by means of visualisation and graphic reports.

Like the analytic review of data, the analysis of data is an iterative process that proceeds one step at a time until a sufficient analytic understanding is reached on the subject under examination.

When the analyses are finished and a sufficient understanding is reached, the results of the analysis are published in the form of reports. Modern technologies of reporting and visualisation enable the work of analysis to continue in the report itself, for instance as drilling into the data.



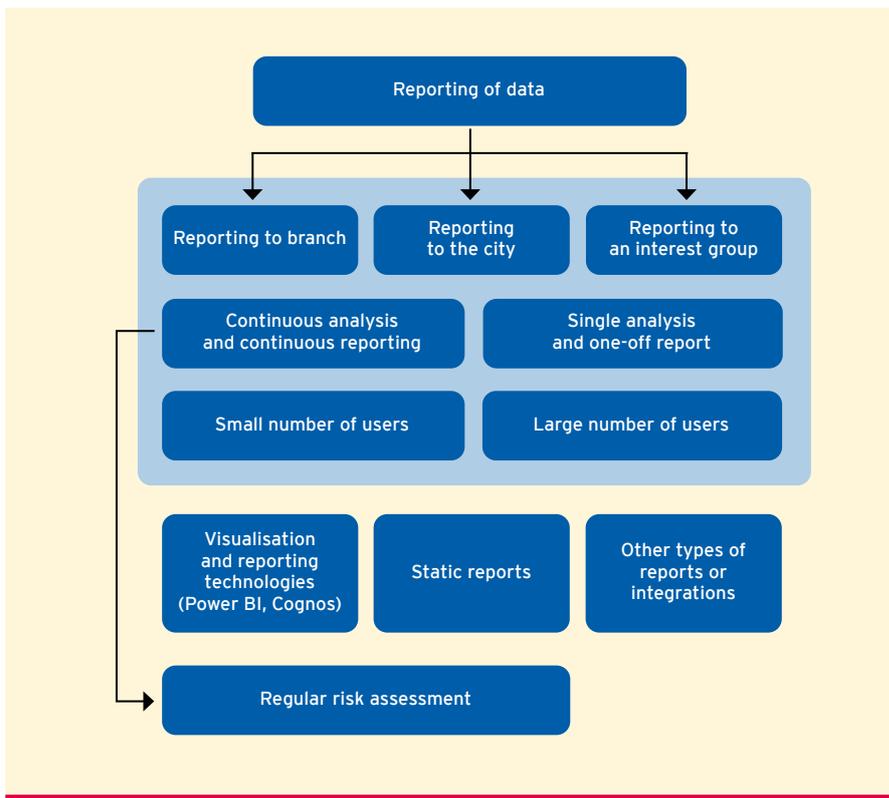
Process description: phase 7 of the permissions process

Reporting and utilising the analyses

Analyses are primarily reported to the party on whose initiative the analysis was performed. The reporting methods depend on who the report is prepared for, but also on whether the analysis and report are intended for continuous use or as a one-off analysis. The choice of method is also affected by whether the report has a small or large number of users.

The object of analysis and the considerations mentioned herein affect the suitability of the various possible ways of reporting the results of analysis. Using reporting technologies is a good choice if the topic of the analysis is many-sided and requires the possibility of drilling down into the data.

Technologies also work well in analyses that require continuous monitoring. If the analysis is one-off in nature or the results need to be delimited more clearly, static reports (e.g. PowerPoint or PDF) are also a good choice. In some cases, there may also be a need to integrate the results of the analysis directly in some other system.



Process description: phase 8 of the permissions process

The chosen means of reporting must be described in the analysis description document and taken into account as part of the risk assessment. For continuous analyses whose risk category is moderate or high, risk assessments must be performed at regular intervals.

The operational model of refining data acts as a kind of handbook that tells how customer information can be used to identify uses for management by data. In this way, services of a higher quality can be developed, among other things. The operational model of refining data provides a basis for efficient utilisation of customer information.

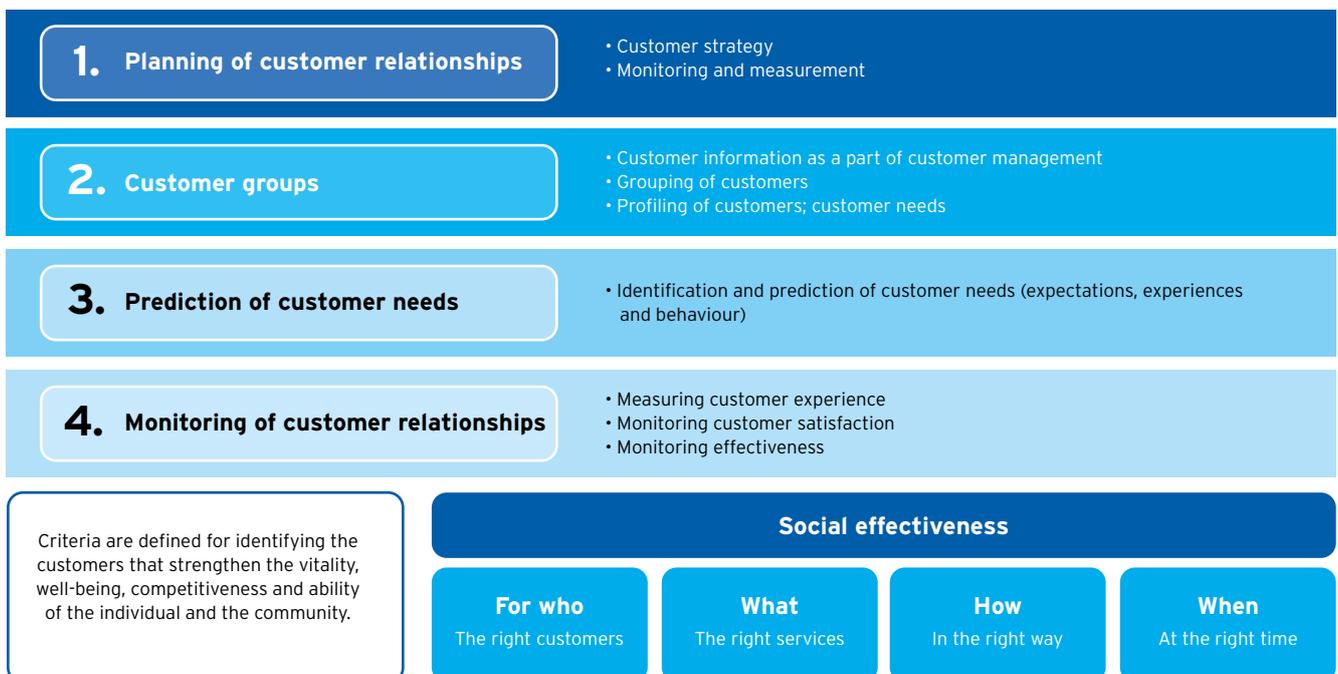
3



Making use of customer information

The “Management of customer relationships” section discusses the planning, grouping, monitoring and measurement of customer relationships with the help of data.

3.1 Management of customer relationships



Planning of customer relationships

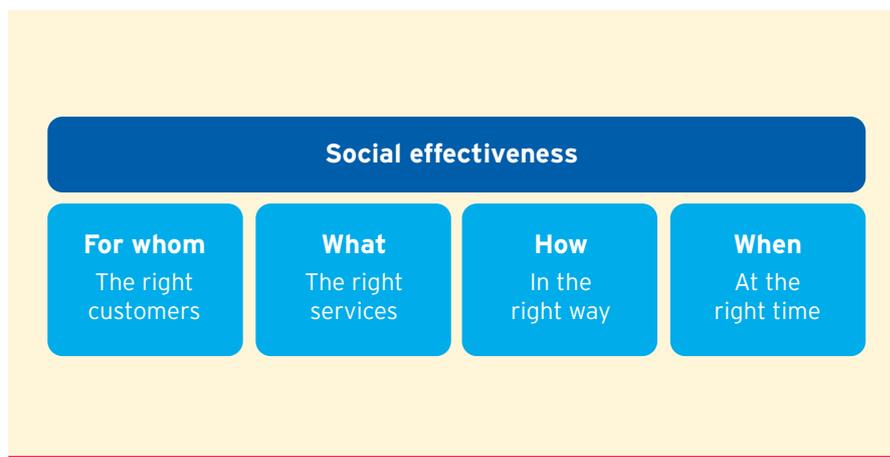
In the planning of customer relationships, a customer relationship strategy is derived from the municipal strategy. The purpose is to steer the development of customer relationships that are successful and beneficial to the city. A model for managing customer relationships and the development of services should be considered in the strategy. The purpose of the customer relationship strategy is also to clarify a shared understanding of what is being aimed at and to set shared indicators for measuring success.

The planning of customer relationships is based on the city's aims and available resources and on information concerning customer needs and volumes. In the planning of customer relationships, criteria are defined for identifying the customers and services concentrating on which will strengthen the vitality, well-being, competitiveness and ability of the individual and the community with the resources available.

Identifying customer and their needs requires the utilisation of information on the operating environment. There are large amounts of information on customers in different data sources (e.g. demographic information, behaviour-related information). With the help of information, the planning of customer relationships and the customer relationship strategy will define:

- **who** the customers are: the customer base of the organisation is analysed and different types of customer relationships are identified
- **what** the customers are offered: the right services are prioritised for the right customer groups
- **how** the customer relationship is handled and the services produced: own processes, tools and interaction with the customer are designed in the right way
- **when** the services and produced to the customers.

In the planning of customer relationships, criteria are defined for identifying the customers who strengthen the vitality, well-being, competitiveness and ability of the individual and the community. In addition, a deep knowledge of customer relationships is also required. Different kinds of information available on customers must be collected from different data sources. For instance, this information may include demographic information or information related to the behaviour of customers. Combining customer information allows customers to be grouped. To enable the right services to be built for customers, it is important to see the big picture.



The parts of the customer relationship strategy

3.1.1 Customer information as a part of customer management

Customer management is based on both the customer and organisation experiencing that they receive value and functional benefit. For this, information is needed on customers and their behaviour. The existence of customer information does not suffice by itself, however. The important thing is how the information is used.

Customer information is collected from several sources and in many different ways, but from only part of all the customer encounters. Neither is the collection and utilisation of customer information usually included in the operative responsibilities at the customer interface. Customer information is gathered, but an overall responsibility for using it is lacking with respect to upper management. In addition to the strategic level, the information should also be utilised in operative working processes.

To ensure the correct functioning of the customer information, it must be integrated and transparent. Integrated customer information means that all the necessary information related to a particular identified customer is connected to the customer in question. The transparency of customer information means that all the organisational levels have one shared picture of the customer. Information related to customer management should be involved in all the functions of the organisation, including the processes of partners and interest groups.

Enabling the utilisation of customer information

To enable customer-centredness to be implemented with the help of information, the systems must genuinely support a customer-oriented business strategy. To enable controlled management of customer relationships, the information received from the systems must be integrated in the processes.

All levels of the organisation - strategic, tactical and operative - must be offered tools, processes and support for customer management. This will ensure the integration and transparency of information.

The multichanneled structure should be managed so that the customer information is the same everywhere, feedback from each customer encounter is integrated to other customer encounters and the customer received the same information and service from each point of encounter. Indicators and analytical tools must also be offered for the management of the customer base.

Getting customers involved is important when the aim is to make the activity as customer-oriented as possible.

Management and utilisation of customer information from the viewpoint of the city

In a digital society, a digital footprint (data) is generated from nearly all interactions with customers.

There are many systems producing data in the city, and their number will probably increase even further. To understand customer behaviour in the future, external data sources or data materials offered by external providers are also needed.

Until now, a large share of the data has been structured, but the share of unstructured data is likely to increase. This development sets new demands for the IT and data architecture, for example from the viewpoints of privacy, data security, data storage and analysis.

On the other hand, the same development enables service development to be "crowdsourced". This means an opportunity for disseminating information to citizens and the city's different interest groups. Programmes related to open data are already in progress, but understanding refined with analytics or the analysis service itself can also be opened for the use of the city's interest groups.

3.1.2 Grouping of customers

Grouping or segmentation of customers can be done in several different ways, depending on the nature of the organisation's activity and the available technology and methods. The grouping should always be constructed carefully on the basis of the need and aim concerned.

Customer groups can be formed and reviewed at both the municipal level and by branch. At the municipal level, customers can be reviewed on the basis of variables such as risk factors and age. Within a branch, the same customer may be grouped on the basis of service use.

Groups can also be formed on several different levels. For instance, we can talk of main groups and subgroups. The main groups may be based on the life phases of the customer, resulting in the following division: child, youth, young adult, adult, middle-aged adult, pensioner. The subgroups would be the result of further divisions within each of these categories. For example, the main group "pensioner" could be divided in subgroups such as actively exercising pensioners and pensioners who attend cultural events.

Regardless of the method used in the grouping, the customer groups to be formed should meet the following requirements:

- The members of the same customer group are similar, e.g. they share the same needs.
- All customer groups should be readily distinguishable from each other.
- Customer groups can be utilised as a basis for a tailored focusing strategy.

The aim of customer grouping is to focus services and marketing on important groups so that they meet the needs of customers. Customer grouping is a means of recognising the needs of customers and of creating services for them.

It would be good to use customer grouping on all the levels of municipal management and development. On the strategic level, the grouping aims to examine the entire lifespan of the customer. On the tactical level, a branch can think of the customer as a user of its own service packages. On the operative level, the grouping can be based on the use of service channels.

The following example describes grouping at different levels:



Grouping of customers on different levels

The phases of the grouping of customers

The phases of the grouping of customers form a whole that the organisation can use to systematically form customer groups and models for managing them that correspond to its own aims.



Customer grouping as a whole

Setting targets

The grouping of customers should be based on the aims of the organisation. Before customers are charted, it is good to clarify the organisation's own aims with respect to customer relationships.

Charting the existing and potential customers and their needs

The needs of existing and potential customers should be discovered with the help of data. Here it is important to know which customers are using the organisation's services. The more detailed information is available, the better.

Charting the factors affecting customer behaviour

The factors affecting customer behaviour can be identified on the basis of customer information and information generated from possible earlier service use. Customers can also be interviewed to enable their behaviour to be better understood.

Grouping is performed, based on suitable criteria

Depending on the organisation's aims, the customers are grouped using the method selected. In grouping, customer information is analysed so that similar factors form clusters that can be used to define the customer groups using suitable criteria. More on grouping methods in "Examples of different grouping methods" is provided below.

Groups are prioritised, if required

When the customer groups are defined, they must be placed in order of priority according to the organisation's aims. The needs of the customer groups having the highest priority determine the services to be developed first.

Creating customer profiles for each group

A customer profile is a more detailed description of typical customers belonging to a customer group. The profile is helpful in designing a model for managing customer relationships.

Aims and benchmarks are planned for each groups

Aims and benchmarks must be designed for each customer group to enable the success of the customer relationship to be measured and monitored.

Development and management model

The customer relationships must be managed and developed continuously. A tailored management model conforming to the aims of the organisation must be designed for each customer group. In addition, customer relationships are monitored and developed continuously.

Examples of different grouping methods

There are several methods for grouping customer relationships. The selection of a suitable method depends on which method is best for guiding the focusing of services and marketing.

The most important methods are presented below.

Method	Description	The information and analyses required
Sociodemographic phase; stage of life	Grouping by demographic factors, e.g. age, gender, training, place of residence, stage of life	<ul style="list-style-type: none"> • Customer information from the population register: age, gender, native language, place of residence • Analysis: lifespan analysis
Based on needs	Grouping by needs, wishes, attitudes, preferences and views	<ul style="list-style-type: none"> • Use of service, phase of life, risk factors, etc. • Analysis of the needs of a customer using the services
Based on behaviour	Grouping on the basis of identified models of service use and behaviour	<ul style="list-style-type: none"> • Number of times the services have been used; other services • Analysis of the users of services by usage (e.g. number of times)
Based on value	Grouping by value of customer (value to city/branch)	<ul style="list-style-type: none"> • Aims of the city, aims of the customer relationship, customer-specific costs, effectiveness • Analysis of the effectiveness of the customer relationship and the use of services
Based on vulnerability	Grouping on the basis of vulnerability and natural inclinations (e.g. at-risk groups)	<ul style="list-style-type: none"> • Risk factor, age • Analysis: another segment (e.g. stage of life) vs. risk

Methods of grouping customer relationships

Information and analyses used in the grouping

The grouping will be more successful the more customer information of high quality is available for use. Below there are some examples of the kinds of information used in grouping:

- Information on the operating environment
- Customer information on the city level
- Customer information by branch
- Customer needs, customer wishes, throughout the lifespan
- External customer information, e.g.
 - Population Information System (VTJ)
 - traffic information system (AKE)
 - national business registers
 - residential area classifications and economic classifications
 - RISC monitors
 - The own databases of analysis businesses
- Stored and analysed customer information
- Qualitative knowledge and tacit knowledge (e.g. the information customer service has on customers, integrated with customer information)

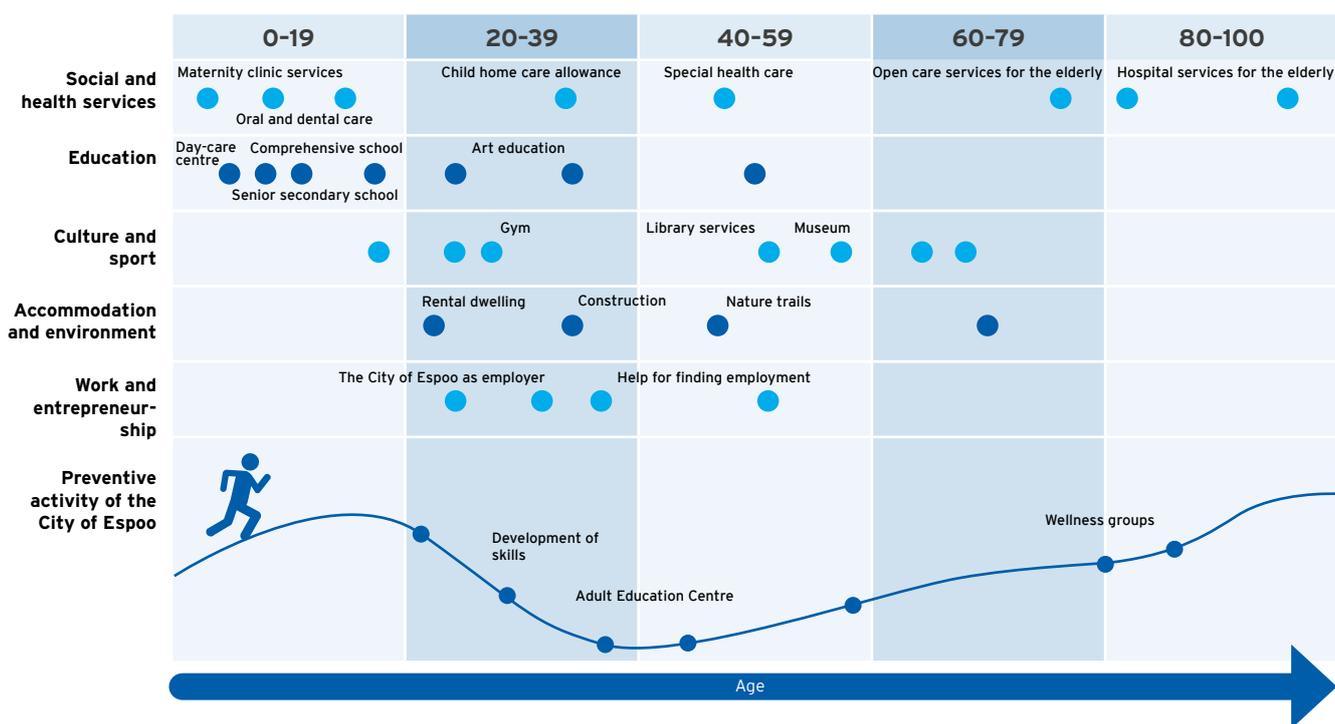
Examples of possible analyses:

- Analyses based on knowledge
- Customer base analysis
- Different tools for segmentation and profiling
- Scoring and potential models
- Value definitions
- Lifespan analyses

Sociodemographic and life-phase-based grouping

A grouping based on the stage of life and the analysis of the lifespan enable the need for services to be predicted.

A grouping using this method enables the age or stage of life of the customers to be taken into account. For instance, the elderly and those in the process of starting a family can form their own groups. The levels of the grouping may vary. For example, a group classified on the basis of age can be divided further according to places of residence, native language and profession.

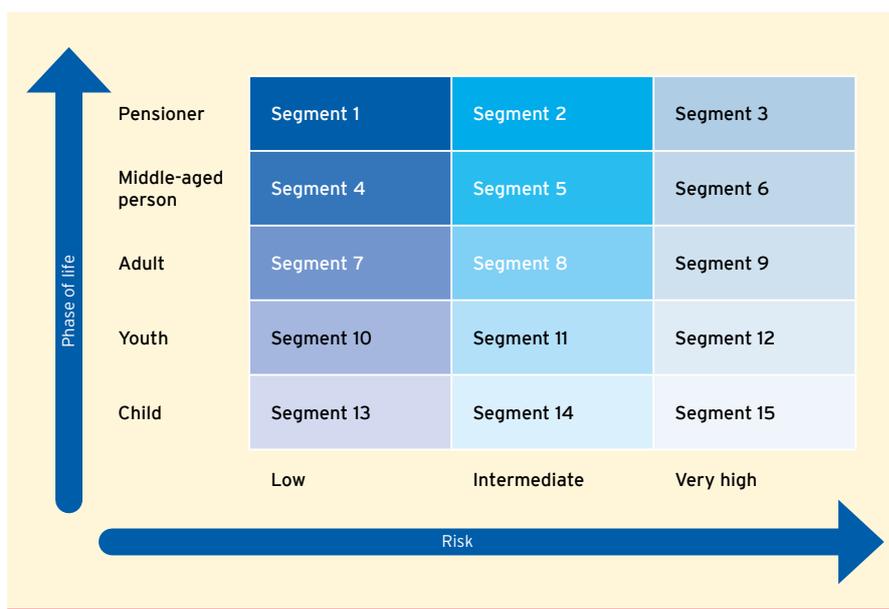


An example of the service needs related to different stages of life

Grouping based on vulnerability

Customer groups can also be examined on the basis of risk factors. Customers in different phases of life are reviewed in terms of their vulnerability to risky situations.

In this method, the customer information to be used may include the customer's age, family situation or income level. This kind of information is available from the Finnish population register.



An example of a grouping based on the phase of life and the vulnerability to risk

Grouping based on needs and behaviour

A segmentation based on needs makes it easier to design customer-oriented services, as the customer group itself is based on the needs experienced by the customers.

A grouping based on behaviour is based on service use. For instance, the group may be determined by customer's level of activity. These kinds of groups include active customers, intermittent customers, customers becoming passive and customers who are already passive.

Customer information related to behaviour may include:

- **customer information** - Who used the service?
- **transaction information** - How and when was the service used?
- **service information** - Which service was used?
- **information on customer service** - Who served the customer and what was the customer offered?
- **the reason for using a service** - Why was the service used?
- **measures directed at the customer** - What kind of communication was there with respect to the situation, and when?
- **feedback information** - Was the customer satisfied with the service?

Grouping based on value

Customer groups can be placed in an order of priority based on how valuable they are to the city in terms of its aims.

The measures of value to be used may include the following:

- profitability, value in euros
- reference value, reputational value
- volume value
- duration
- strategic value
- potential value (prevention)

Describing the customer group and customer profile

When the customer information is analysed and the customer groups have been formed, e.g. with the help of clustering, the groups should be described in a clear way. It would be good to prepare the description in the same way for all groups, making it easy to identify the differences between groups. For example, the description may include a basic description of the group and its most typical identifying elements, factors affecting behaviour, customer needs and the service producer's promise of value to the customer.

To enable services, marketing and activity to be developed, the descriptions prepared on customer groups should be distributed to everyone within the organisation. All members of the organisation should have a shared view of the different groups and their needs.

Name of customer group	
Description <ul style="list-style-type: none"> • What are the elements of identification of the group; how does the group in question differ from other groups? 	Factors affecting behaviour <ul style="list-style-type: none"> • Which factors have the biggest influence on the behaviour of the customer group?
Customer needs <ul style="list-style-type: none"> • What are the primary needs of the customers in the group? • The service needs to be responded to? 	Promise of value <ul style="list-style-type: none"> • How can the city/branch/service support the customer group in question?
The criticality/prioritisation of groups <ul style="list-style-type: none"> • How significant is the group in question? Is its priority high, and if so, why? • Compatibility with the strategy 	

An example of a tool used in describing a customer group

The customer group description is often a more upper-level description of first-level groupings. Within the same customer group, there may nevertheless exist several different profiles; it is useful to describe them in more detail. For example, the customer group may be "students", but a student may have one of several different profiles (e.g. student at primary or secondary school). For this reason, it is advisable to describe the customer profile on the level on which the services and marketing are focused.

The customer profile is a description of the archetypes of services. It can be used to identify the customer and the way in which the customer uses a service. In addition, the customer profile contains a description of what is valued by the type of customer in question. In addition to tailoring services, profiles can also be used in developing customer relationships and in measuring the effectiveness and customer satisfaction of the service being offered.

Customer	
Basic information	<ul style="list-style-type: none"> • Basic information on the customer • Key background information, e.g. training, hobbies
Demographic information	<ul style="list-style-type: none"> • Gender • Age • Wealth class • Place of residence
Identification	<ul style="list-style-type: none"> • Mannerisms
Aims	<ul style="list-style-type: none"> • Key aim • Other aims
Challenges	<ul style="list-style-type: none"> • What is the biggest challenge of the customer? • What other challenges does the customer have?
Promise of value	<ul style="list-style-type: none"> • How will the customer be helped with their problem? • How is the customer helped to reach their goal?
The most important message	<ul style="list-style-type: none"> • How is the service described/communicated to the customer?
Primary channels	<ul style="list-style-type: none"> • Which channels work best?

Customer profile

CASE ESPOO

Museum

During the project, descriptions of customer groups were prepared for the Espoo City Museum. The descriptions improved the understanding of customer behaviour and helped to develop the services of the museum to better meet the needs of customers. As one result, the museum can plan focused marketing

and events for different customer groups while emphasising the factors that make it stand out. In addition, the results of the work will enable the museum to plan its future activities in a more customer-oriented way.

The customer groups that were identified needed information, conviviality, sociality, activities, the company of peers, well-being and advice.

Learners and people interested in knowledge



Need:

- Learning new things and collecting information

Leisure-time users, visitors and travellers



Need:

- Trendy things to do
- Experiences
- Locations to photograph for social media
- Sociality

Persons needing service from officials



Need:

- Statement
- Advice
- Support for communications

Life-phase-based visit/ those searching for a peer group



Need:

- Activities for children/ /the elderly /parents
- Meeting other members of the peer group
- Becoming more cultured

Employees and associations



Need:

- Meeting work-related responsibilities (e.g. in-school curriculum)
- Support for work
- Interest

Those coming via social and health services



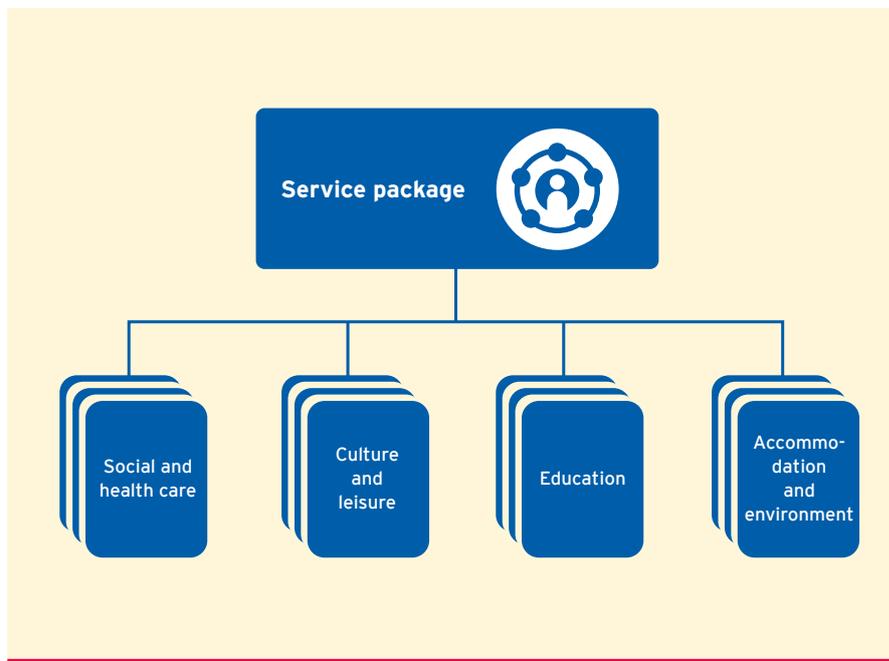
Need:

- Well-being
- Learning
- Sociality

The identified needs of customer groups in the Espoo City Museum

3.1.3 Prediction of customer needs

With the help of analytics, the needs of grouped customers can be predicted, and the right service packages can be offered to customers in good time. The service packages can be offered across branch boundaries through own service production or by directing the customer to a suitable outside service.



A service package is formed from the services of different branches, based on the predicted customer needs

Prediction of customer needs (expectations, experiences and behaviour) can have many kinds of benefits. A view of the need for services can be formed by analysing information on customers and the operating environment. Principles can be created for constructing service packages; service plans can be prepared for customer groups; services can be prioritised as a part of customer management; and services that meet or exceed customer expectations can be offered easier than before.

For example, predicting how the service paths are determined is made possible by finding out how starting to use a service affects the use of other services. Finding out correlations requires combining the customer information of different services. When a possible correlation is sought between the use of different services, the following questions among others should be answered:

- How much are the municipal services examined being used?
- What is the user profile of these services?
- Which municipal services are used before and after a particular service that is being examined?
- What can be the indicator for starting to use a service?

When preparing predictions, all customer information should be gathered in one database that is then used as a source for predictive analysis (here, correlation analysis).

Among other things, the following pieces of information can be used in the analysis of predicting customer needs:

- Information on the operational environment
- Basic customer information at the municipal level
- Customer information by field
- Information received from interest groups
- Customer needs (information on customers' wishes) throughout the lifespan
- Service information
- Service volume
- Capacity information
- Trend information

3.1.4 Monitoring of customer relationships

By monitoring customer relationships, a comprehensive view can be formed of the effectiveness of the city's activity throughout the service units. The purpose of the monitoring is to make the management of customer relationships as transparent as possible.

Monitoring and measurements are carried out on three different levels:

Strategic level

On the strategic level, changes in the lifespan of the customers are monitored, together with the effectiveness of the measures taken. Examples of customer-oriented yardsticks:

- Changes in customers' expectations, experiences and behaviour
- Duration of the customer relationship
- Development of the customer relationship
- Qualities of the customer
- Customer portfolio
- Efficiency of the services
- Customer profitability
- Share of new customers

Tactical level

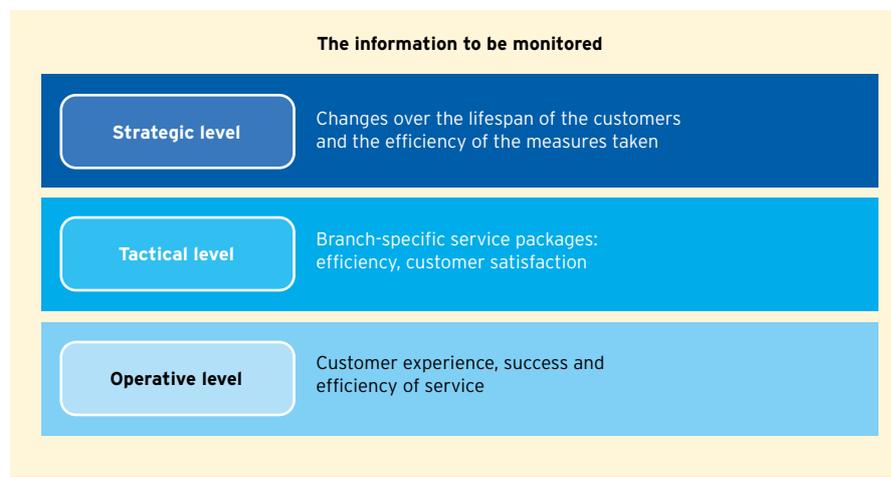
The effectiveness of branch-specific service packages is followed on the tactical level, together with customer satisfaction. Examples of customer-oriented yardsticks:

- Changes in customers' expectations, experiences and behaviour
- Indicators related to service channels
- Efficiency of the services
- Qualities of the customer
- Service visits by the customer
- Customer profitability
- Share of new customers

Operative level

The operative level monitors customer experience and the success and effectiveness of the service. Examples of customer-oriented yardsticks:

- Changes in customers' expectations, experiences and behaviour
- Customer activities
- Efficiency of the services
- Service visits by customer

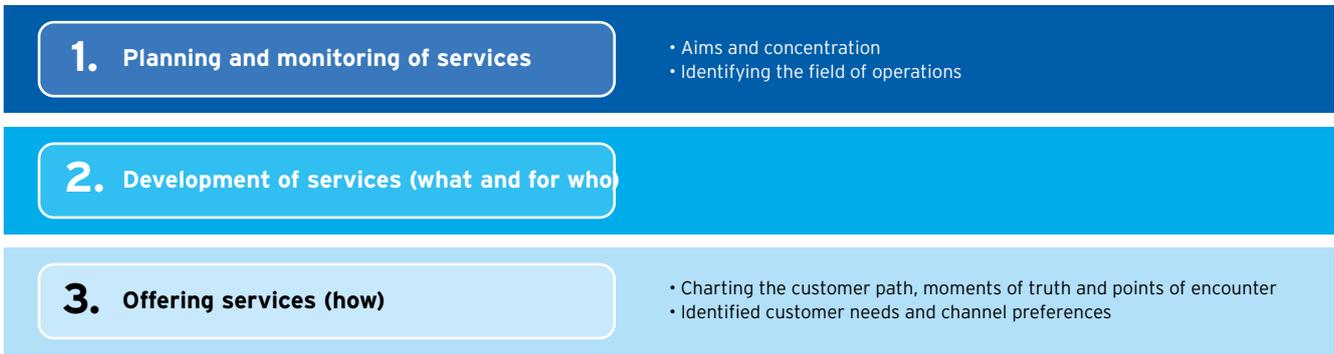


Monitoring of customer relationships

3.2 Arranging services

Arranging services means arranging and coordinating the service offering, the service channels and the competences. This also includes the more detailed definition of the municipal offering and services and the selection of service producers. The services are arranged or acquired from internal or external producers on the basis of service needs; the preconditions of customer management; the strategy and financial management. The production contracts are administered and the services themselves are ordered from their producers.

In terms of information, the important thing here is to analyse the effectiveness of the current service offering with the help of data, to utilise customer information in developing the service offering and to monitor the effectiveness of the services with data.



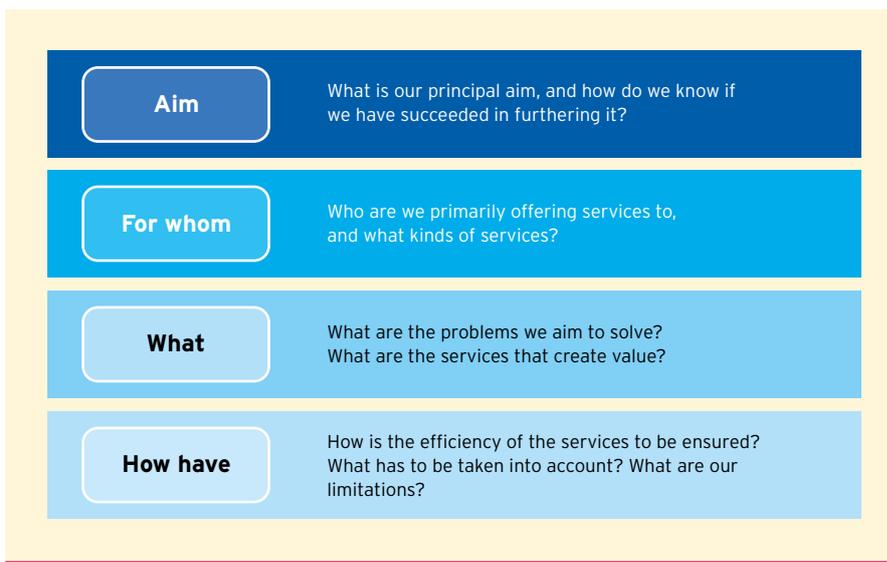
Arranging services

3.2.1 Planning and monitoring services

When services are arranged, it is crucially important to define the organisation's service offering. The service offering is defined by identifying the services that are currently being offered and the services that should be offered in order to meet the set aims.

The starting points for the service offering are the strategy and aims. The mission indicates who the organisation exists for: who should the services be directed at on the upper level? The vision helps to define the direction of service development based on where the organisation is going. In addition, the added value generated by the organisation for the customer in the present and the future must also be defined.

The key questions of service design, which should be answered when the service offering is being designed and developed, are presented below.



Planning and monitoring of services

Aims of the city and the service units

The city designs the upper-level aims and strategy, which offer guidelines for the development of services. The city's customers should be monitored throughout all the service areas. The lifespan of all the residents of the city and the related services should be examined. If this is done, the city is able to steer the whole to ensure that the right residents receive the right services at the right time. Steering means indicating a direction based on knowledge. This is done by enabling the proactive offering of services by service units in the future.

The service unit in turn carries out its own aims, often limiting its examination to its own customer relationships. Service units should also examine customer relationships in a larger context than just its own service. The service provided by a service unit may be significant to other events over a customer's lifespan, and thus also to the realisation of the city's aims. Although the aims of service units are focused on the units themselves, the overall picture on the strategic level should always be kept in mind when preparing a service offering.

Identifying the operating environment

Assessment of the operating environment is important to the planning of a service offering. Through the assessment work, a view is formed concerning the environment of acting, what is intended to be achieved and what is wanted to be changed. Information on the operating environment offers a view of the volumes required and the development of service needs. To enable the right services to be developed, there should be a versatile and comprehensive knowledge of the field of operations.

Among other things, the field of operations means the environment related to politics, nature, consumer behaviour, technology and law. The analysis to be carried out is called a PESTEL analysis. Its purpose is to bring into view the effects that the changes related to the operating environment will have on customer relationships and service needs.

Information on the operating environment supports decision-making by the service provider when it wished to prepare for, and react to, changes in the environment. For instance, changes to the legislation may affect the way in which services must be offered in the future, even in cases where the customers would wish otherwise.

The results of the assessment should always be combined with information collected about customers, partners and other businesses. In this way, it can be seen as to whether the service offering should be expanded by developing entirely new services or already existing ones.

3.2.2 Service development

This section describes the methods and tools of product development and commodification.

The service offering of a business must develop along with the needs of its customers and operating environment. To enable services to be developed continuously, the service offering should be reassessed at regular intervals. New ideas for services must always fit in with the wider strategic aims.

In this context, service development means the development of both new and existing services. The levels of service development can be divided in gradual and radical innovations, depending on whether the current offering is being developed or something new is being created. In addition, the division may be related to whether the need being met is an existing one or a wholly new one. In the framework at hand, the process of development is the same regardless of whether the service being developed is a new one or an already existing one.

1. Changing an existing service	<ul style="list-style-type: none"> • Improvements to an existing service, e.g. in relation to its features or service production 	} Gradual innovations
2. Extension of the current service offering	<ul style="list-style-type: none"> • Expanding a service, e.g. by offering new content • Adding new areas to an existing service 	
3. A new service to meet an existing need	<ul style="list-style-type: none"> • A new service to meet an existing need, which up to now has been met in some other way (e.g. another service) 	} Radical innovations
4. A new service to meet a new need	<ul style="list-style-type: none"> • A new service to meet a newly identified need that has not been met before • For instance, the needs may have been identified by analysing data or through customer feedback 	

The levels of service development can be divided in gradual and radical innovations

The levels of service development

The process of service development

The process of service development describes the phases of service development from identifying the need for a service to developing a service package, taking it in use and developing it further.



The process of service development

Clarifying the aims

The process of service development starts from clarifying the aims. Before more detailed plans are made for meeting customer needs, the needs of the organisation have to be made clear. The aims should describe clearly what the services are aiming at in the first place.

During this phase, it is good to go through the current service offering and compare it to the aim. This will bring into view any possible gaps or deficiencies.

When the need for a service is clear, its strategic suitability and implementability can be tested, for example through the following questions:

- Does the service carry out the strategy of the organisation?
- Does the service add to the current service offering of the organisation or correspond to it?
- What kinds of competences and resources does the implementation of the service require? What kind of resources are available to use?
- What kinds of risks are related to the service? What kind of effect would the service have in terms of growth and competitiveness?

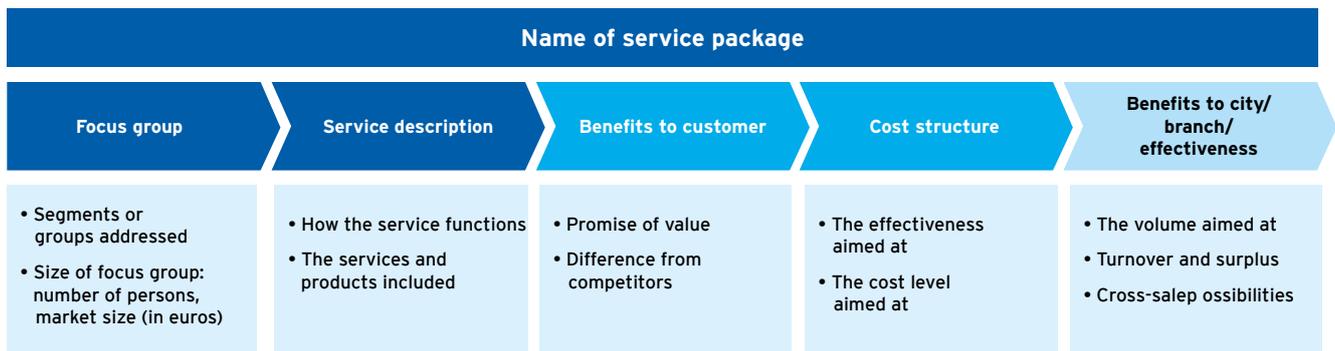
Understanding customers and their needs

Understanding of customers means understanding the added value created by customers and one's own service. Understanding of customers can be reached by refining customer information. Customer information can also be utilised as a support for business decisions.

Customer-oriented service development requires sufficient information on the needs and challenges of the customer. Understanding of customers is needed when new or existing services are being developed. It can also be used to find out the market for a new service or one that is being developed.

In a service, the customers themselves participate in the formation of value, so it is important to have a better understanding of how they would like the service to be implemented. Customers should be involved in service development by taking them along in development workshops or interviewing them.

Planning service packages



Description of the service concept

When the need for a service has been recognised and found to correspond to the organisation's own strategy and the customer groups having the need have been identified, it must be considered how the service should be implemented as a whole. This section concentrates on presenting a method and a tool for planning a service package as a whole.

When planning a service package, use can be made of Kuntakanvas (kuntakanvas.fi), a strategic tool for structuring and managing activity. Kuntakanvas can be used to grasp an activity as a whole and to assess the changes occurring in it with respect to customer relationships, competences and finances.

Kuntakanvas consists of two main levels and 11 areas, which can be used to describe a service package while taking into account the strategic aims and the promise of value given to the customers. With Kuntakanvas, service activity can be planned from the present state to the target state.

In the use of Kuntakanvas, it is intended to involve personnel and interest groups in developing activity together. Kuntakanvas helps to steer thinking about activity in a customer-oriented direction. It asks: Who are all the people we are producing services and added value for?

It is recommended to fill in Kuntakanvas in the following order:

Customer relationship and the promise of value

First it is important to think about customer needs and customers. On the basis of this, a promise of value based on the municipal strategy can be filled in. The promise of value is a description of the benefit offered by a service from a defined customer viewpoint. Realisation of the promise of value should be considered separately for each customer group.

With respect to customer relationships and the promise of value, key questions include the following:

- Who is the service intended for?
- What kind of benefit are customers aiming at when purchasing the service?
- How many potential customers are there?

Service offering, channels, laws and competences

Under consideration next are the services to be offered, and the service points and channels through which value is to be produced to different customer groups. All the relevant legislation and other key recommendations governing the arrangement of services must also be taken into account.

In the service description, the structure of the service is defined, together with commodification (core, support and supplementary services, plus modulars). In this phase, those parts of the service are described that are produced as standardised for all customers, together with those that are tailored for different customers. If actors other than the organisation itself will participate in producing the service, it is advisable to involve the possible partners in the work of definition.

Competences are all the factors that are required for a successful service that brings value to customers. Competences consist of models of acting, processes, information, channels, funding, expertise and resources. Resources are persons, materials, facilities, information systems, data repositories, technology, tools, permissions, rights and agreements.

Key processes and interest groups

It is described as to how the services are produced, managed and monitored. Which interest groups and partners are needed to help with producing and developing the service? Here it is important to consider whether the service will be arranged together with customers, a centralised unit of the city or external actors.

The core processes of the activity are identified. These include the processes related to customer management and the provision and production of services. The key processes may be support processes, such as personnel and financial management, or steering processes, such as strategic management or management by results.

Cost structure and income streams

When the infrastructure of arranging an activity or service is sketched out, the question of financial balance is considered: what do the costs of the activity or service consist of, and how will it be funded?

Here it is useful to assess the market potential of the service and the future sales volume. For instance, the market potential can be identified by analysing population-related data.

Effectiveness

In effectiveness, it is again a question of customer relationships: what are the effects of our activity? Different viewpoints must be considered here, but especially that of the effect on the customer. What are the measurable results, and what is used to measure them? The service package can be described with Kuntakanvas.

Customer relationships and services	CUSTOMER RELATIONSHIPS 	VALUE PROMISE 	SERVICE OFFERING 	KEY PROCESSES 
		LAWS 	SERVICE CHANNELS 	STAKEHOLDERS: 
Financial balance	COST STRUCTURE 	REVENUE STREAM 	CAPABILITIES 	EFFICIENCY 

The Kuntakanvas tool for developing service packages

Defining the process descriptions and introduction of the services

When the service package is sketched out with Kuntakanvas, a process description is prepared on the service. This describes how the service is produced and implemented.

The customer is usually involved in the production process of a service. The service process covers functions related to the production of services both within the organisation and at the customer interface. Part of the process is visible to the customer and part invisible.

It pays to describe the phases of implementing the service in as much detail as possible. Describing the service process is helpful in defining who participates in the production of the service, and in which phases.

In addition, the introduction of a service involves training on the content of the service and the service process to the personnel responsible for service production and to other interest groups.

Monitoring, assessing and developing the services

It is important to monitor how the services meet the aims of the city and the service unit.

The key indicators of success in the management of services are:

- effectiveness
- cost-effectiveness
- customer experience

In addition, the quality of the service must be monitored. A quality target must be defined for services to enable them to be monitored continuously. The quality of a service can be considered good when the customers' experience of the service corresponds to their expectations. The customer experience is affected by the expected and experienced quality of the service.

It is easier to develop services continuously and systematically when they are monitored in a continuous and systematic manner. Customer information is an important part of monitoring and assessment. Customer information should be monitored, assessed and predicted across the boundaries of service units.

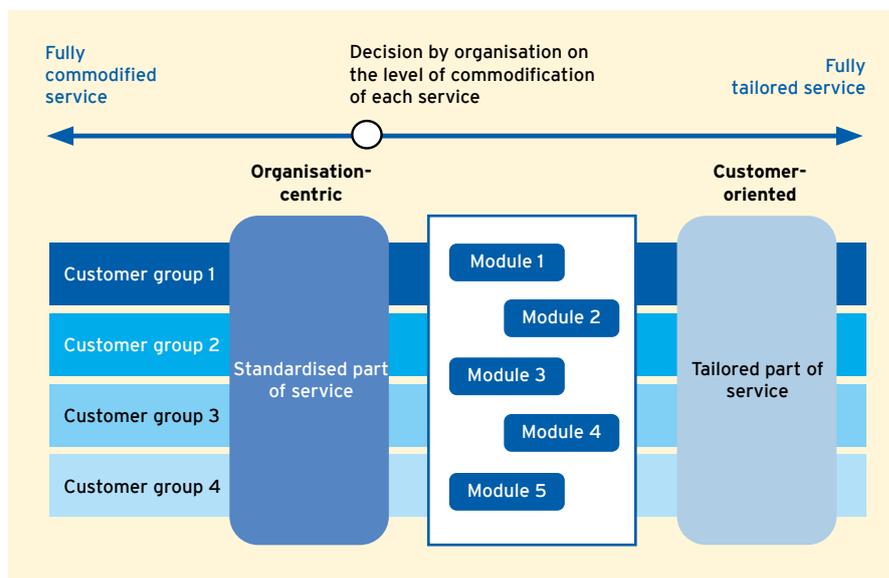
3.2.3 Commodification and tailoring of services

Services can be produced as standardised or tailored services or in a modular way. Building a service as completely standardised leaves little room for the individual needs of customers. On the other hand, it is not reasonable to build all the services from scratch for each separate customer.

With commodification, part of a service can be standardised. The service does not have to be redeveloped each time; it is instead offered to each customer in the same form. Commodification reduces the need for tailoring to a reasonable level.

In addition, the parts of the service (modules) can be tailored for each customer with different needs. To some extent, tailoring always belongs to the basic nature of the service. Commodification and tailoring support each other and form a functional service package. An intermediate form between standardised and tailored is offering a service partly or wholly in modular form. Customers can then select the parts of the service they need.

The final offering of services should be thought about separately for each customer group.



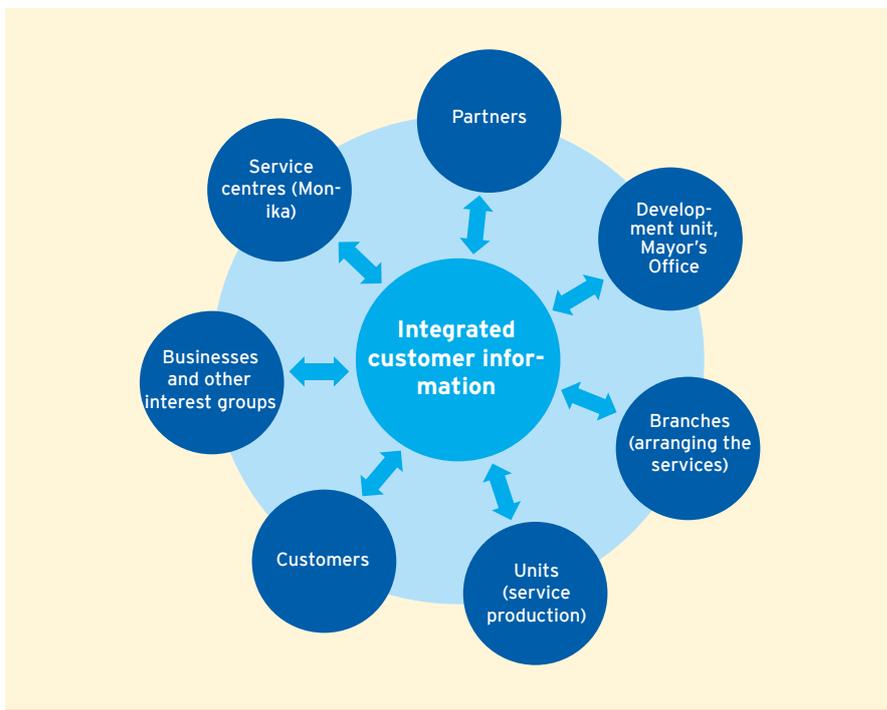
Commodification and tailoring of services - modular service structure (Sipilä 1995)

3.2.4 Managing the development of activity

With respect to service development, the city acts as an enabler. To its service units and interest groups, the city offers a platform for efficient, continuous and customer-oriented development of services according to the city's own aims.

The purpose of the platform is to collect, disseminate and produce as much information as possible with respect to customers and services. Pieces of information received from interest groups can be analysed in one place and integrated with each other. The information can then be offered onward for the use of interest groups. With the help of the information, customer relationships and services can be developed to meet the changes occurring in the operating environment.

The processes and tools used by the organisation and the interest groups should facilitate the integration and utilisation of information.



Management of the development of activity

Offering services

Offering services can be based on a need for services or on predicting such a need. Making use of customer information and predicting needs is described in more detail in section 1.3.

The customer path and the points of encounter related to it are very significant in the development of services. Charting the customer path is a precondition for developing the customer experience. The purpose is to find out a customer's important experiences and points of encounter during the service. For each point of encounter, knowledge and understanding should be collected with respect to their effect on the customer. This knowledge can be used when thinking about how services could be offered even better in the future.

The customer path also provides information about the identified customer needs and channel preferences. During the customer path, both the customer's own activity and the service producer's visible and invisible activity are gone through,

3.3 Service production

1. Services for selected customers

- Utilisation of customer information
- Service levels according to customer group

2. Management and monitoring of service production

- Utilisation of customer information
- Production of customer information
- Monitoring the customer path

Service production

3.3.1 Utilising customer information

By customer management and by making use of customer data, services can be better focused on different customer segments in encounters with customers.

The information available for use in encounters with customers includes the following:

- wide-scale customer information, i.e. lifespan information and analyses
- information on service needs, i.e. information based on history
- information based on customer feedback

The uses of customer information in service production are the following:

- preventative guidance
 - e.g. the child protection flagging (notification) used by the City of Espoo: could this be combined with information on the customers of addictions services?
- recommendation or marketing of services based on customer segment, profile and service history, e.g. book recommendations by the city library or recommendations related to culture and sport
- developing the service process
- developing the customer relationship
- collecting customer feedback

3.3.2 Service levels

By systematic customer management and by making use of customer data, segment-specific services can be implemented and service levels, indicators and aims can be defined for them.

The following table shows how a set of indicators can be built by customer group.

	Indicators	Aim	Current situation
Customer group 1	Indicator 1		
	Indicator 2		
	Indicator 3		
	Indicator 4		
Customer group 2	Indicator 1		
	Indicator 2		
	Indicator 3		
	Indicator 4		

Service levels

3.3.3 Management and monitoring of service production

Service production is the point where the customer is encountered. Customer data is both utilised and created there. As a rule, each encounter with a customer produces new knowledge.

Among other things, the knowledge produced in service production includes the following:

- knowledge related to use of services by customers
- knowledge based on customer feedback
- knowledge based on customer surveys
- knowledge produced by customers with their own devices

Through the systems, the knowledge is returned to the process of analysis, enabling the development of measures, estimating their effectiveness and facilitating future development.

The monitoring and analysis of the customer data passing through service production helps to measure:

- the direct effects of services
- the effectiveness of services in the long term
- the functionality of developmental measures
- the functionality of service production by third parties
- customer satisfaction
- the functionality of a customer strategy

The data passing through service production is both a yardstick for the success of an activity and a "raw material" in the planning and analysis of new objects of development.

Concepts

Analysis. A result or conclusion produced by the process of analysing data

Customer. An active actor, whether an individual, a business, a community, society or environment, can be a potential new customer or an existing one, having a unique logic to their activity

Customer experience. The customer's subjective experience of being a customer

Customer-centredness. In the values, strategy, quality policy and everyday activity of the organisation, a way of thinking and acting that starts from the needs, interests and rights of the customer. In customer-centred service, the needs, wishes and expectations of the customer are taken into account without interpreting them, as the customer expresses them.

Customer service. An encounter between a customer and a service organisation on some channel.

Customer group. A group of customers united by similarities (e.g. age). Customer groups can be formed in many different ways, depending on need and the benefit aimed at.

Understanding of customers. Identifying customers' wishes in customer encounters in a situation-dependent way. Enables the identification of the need for services.

Management of customer relationships and services. Management and utilisation of the emphasis on customers. Aims at ensuring the practical realisation of the emphasis on customers and requires systemic management of competences.

Customer management. Management of customer relationships in a goal-oriented and structured way. Management and utilisation of customer information and the understanding of customers have a central role here.

Customer relationship. A relationship based on a partnership between the customer and service provider. The scope varies from a possible new relationship to existing relationships whose scope may differ.

Customer information. Information on the experience of being a customer, on using services and on the aims of customer relationships. The information can be "raw data" taken from systems or more refined information. The information guides the planning and development of activity.

Data. Unrefined, transaction-level "raw data" in repositories or other information systems, often in electronic form.

Service management. Focusing of competences with emphasis on the customer. Aims at focusing resources in an effective and cost-effective way, with systemic and agile management of resources as the starting point.

Service provision. Defining and arranging municipal services; selection of suppliers in cases where service production is outsourced.

Service production. Producing the services commissioned by a party responsible for arranging municipal services, according to the quality targets.

Information. Raw data refined through increasing one's understanding of it. For instance, data is turned into information when raw data is grouped.

Refining data. A process of refining data into information.

Analysis of data. Often a synonym for refining data, if the refining is done with analytic methods. "Refining data" and "analysing data" are used interchangeably in this document

Appendix 1

Benefit, risk and cost estimate form for analysed data

1. Basic information	61
2. Summary of the risk/benefit/cost estimates	62
3. Potential benefit of analysis	63
4. Potential risk related to data analysis and reporting	65
5. Costs related to data processing and analysis	68
6. Overall estimate	69

1. Basic information

Basic information on the object of analysis

Analysis	
Name of analysis:	
Description of analysis:	(A short description of the topic analysed)
Analysis commissioned by:	(Field, responsible person within the field)
Analyst:	(The person analysing and refining the data)
Data retrieved by	(The person(s) who fetch the data from the source system and anonymise it if required)
Purpose of the analysis:	Internal/External
Users of the analysis	
Means of publication of the analysis	(How are the end results of the analysis presented?)
Continuity of the analysis	Continuous monitoring, updated at intervals, one-off
Data	
Content of data:	General description
Customer data:	Contains customer data Yes/No
Time frame of data	Time frame used in analyses
Source(s) of the data:	(In what system is the data located?)
Privacy protection of data:	Data anonymisation methods
Data security	Data fetching and transmission methods
Data processing tools	Information systems
Owner of the data:	(Field, responsible person within the field)

2. Summary of the risk/benefit/cost estimates

Preliminary assessment

Risk assessment: Low Moderate High

Benefit: Low Moderate High

Cost: Low Moderate High

Initiator of the analysis: briefly describe your view of the risks, benefits and costs of the analysis

Write here...

Statement of the inspection group

Write here...

3. Potential benefit of analysis

- The potential benefit of the data analysed must be identified, and the amount of benefit must be assessed.
- In many respects, a detailed picture of the potential benefits is impossible to achieve, but the purpose of this section is to create a preliminary assessment that can be improved later on.
- In the overall assessment, the assessment of benefit is assessed in relation to the risks identified and the costs caused by processing and analysing the data.
- A low potential benefit does not mean that an analysis cannot be performed.
- The assessment of benefit helps the parties responsible for analyses to place the objects of analysis into an order of priority.

Describe the parties that are likely to make use of the analyses and their results.

- Different units of the City of Espoo
- Other actors in the public sector
- Residents/citizens
- Businesses
- Non-profit organisations
- Research institutions or researchers
- The media
- Other parties, which? _____

Describe briefly what the analyses aim at and how they will be utilised

Write a short description of the aims of, and ways of utilising, the analyses here...

Use the table below to assess the overall benefit of the data to be opened.
Consider:

- A. The benefit of the analysed data to its users (x axis)
- B. Probability that the analysed data will be utilised/used (y axis)

Probability of using the open data	Very high probability	Slight benefit	Moderate benefit	Great benefit	Very great benefit	Very great benefit
	High probability	Slight benefit	Moderate benefit	Moderate benefit	Great benefit	Very great benefit
	Moderate probability	Slight benefit	Slight benefit	Moderate benefit	Moderate benefit	Great benefit
	Low probability	Very slight benefit	Slight benefit	Slight benefit	Moderate benefit	Moderate benefit
	Very low probability	Very slight benefit	Very slight benefit	Slight benefit	Slight benefit	Slight benefit
		Very slight significance	Slight significance	Moderate significance	Great significance	Very great significance
		Estimated potential significance of the open data to the users identified				

Give your overall estimate of the potential benefit of the data to be published

- Very slight benefit = 0 points
- Slight benefit = 2 points
- Moderate benefit = 5 points
- Great benefit = 8 points
- Very great benefit = 10 points

The overall estimate is:

4. Potential risk related to data analysis and reporting

- With respect to the analysed (and combined) data, risks related to the content of the data and the publication of the data must be identified.
- This form is intended to assess risks on two levels.
- On the first level, the identified risks are related to personal information and the requirements set by the new GDPR regulation.
- On the second level, other possible risks are identified. These may be related to the possible reputational risk created by the reported data, citizen reactions to opening up the data, legal risks, general safety or other possible data-specific risks.

Personal data

- The current legal definition of personal data is in Section 3(1) of the Finnish Personal Data Act: personal data means any information on a private individual and any information on their personal characteristics or personal circumstances, where these are identifiable as concerning them or the members of their family or household.
- Thus, personal data means all the data which can be used to identify a natural person, i.e. a human being. Among other things, this data includes a person's name, personal number, address, e-mail address and photograph.
- The GDPR came into force on 25 May 2018, and it gives the definition of personal data.
- The definition is in Article 4(1) of the GDPR: "personal data" means any information relating to an identified or identifiable natural person ("data subject"); an identifiable natural person is one who **can be identified, directly or indirectly**, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.
- According to the GDPR, personal data thus means all data that can be used to identify a natural person (a human being) directly or indirectly. Among other things, this data includes a person's name, telephone number, personal number and photograph.

Assess whether the so-called "raw data" includes personal data

- Yes
- No

Assess whether the data analysed includes personal data

- Yes
- No
- Unable to assess (further assessment necessary)

Assess whether the end results of the analyses (e.g. reports) include personal data

- Yes
- No
- Unable to assess (further assessment necessary)

-
- If the analysed data includes personal data, but the end results of the analyses do not require personal data, **the analysed data must be anonymised** as well as possible so that natural persons and their data cannot be identified directly or indirectly.
 - In spite of anonymisation, there is always a risk of natural persons or their data being identified. For instance, in the future, this risk can be realised by combining different pieces of data and the information they contain and thus creating a new means of identifying the information in the original data. Future technological developments (e.g. artificial intelligence) may also enable the use of more efficient statistical algorithms, increasing the risk of identification.
 - If the analysed data includes personal data that has not been anonymised, the risk of the analysis is always assessed as “high” or “very high”.

Assess and indicate what kind of harm may be caused by the identification of the personal data contained in the data

- Significant harm (e.g. identity theft, discrimination, abuse) may be caused to natural persons by the dissemination of personal data contained in the data.
- Dissemination of information contained in the data may reveal location information that can cause harm to natural persons (e.g. break-in, theft, damage to property).
- Dissemination of the information contained in the data can cause economic harm or loss of earnings to natural persons.
- Dissemination of the information contained in the data can cause psychological harm to natural persons (e.g. loss of reputation, attacks on reputation).
- Some other harm, which? _____

Other risk factors related to data content

- Even if the data in the process of analysis does not contain data that can be considered personal, publishing the data can **also involve other risks**.
- For instance, an excessively detailed publication of the location and time of parking fines given by municipal traffic wardens may cause a risk of identifying the daily routes of individual traffic wardens. This may in turn cause an increased risk of physical or psychological violence against the traffic wardens.
- Opening up or revealing the information included in the data may also involve risks to the reputation, juridical position or operative activity of the City of Espoo.
- Although the published data is anonymised as well as possible or does not contain any personal information in the first place, public opinion may still view its publication as inappropriate.

Assess and record what other kinds of damage the publication of the information in the data may cause

- Publishing the data or revealing information contained in it may cause significant damage to the public image of the City of Espoo.
- Publishing the data or revealing information contained in it may weaken the operative activity of the various offices or branches of the City of Espoo.
- The data contains data given to the city by a third party, for whose publication there is no separate permission
- The data contains detailed information that may weaken general safety (cf. e.g. too detailed architectural drawings of public buildings)
- Publishing the data or revealing information contained in it may cause significant harm to the City of Espoo in that city residents are no longer willing to provide all the information required for managing the city.
- Publishing the data is likely to be experienced as inappropriate by citizens

Overall estimate of the overall risk of the analysis

Use the table below to assess the overall risk related to the analysis to be undertaken. Consider:

- A. The potential harm related to the analysed and reported data (x axis)
- B. The probability that the risk of harm will be realised (y axis)

Probability of realisation of the harm related to the open data	Very high probability	Low risk	Moderate risk	High risk	Very high risk	Very high risk
	High probability	Low risk	Moderate risk	Moderate risk	High risk	Very high risk
	Moderate probability	Low risk	Low risk	Moderate risk	Moderate risk	High risk
	Low probability	Very low risk	Low risk	Low risk	Moderate risk	Moderate risk
	Very low probability	Very low risk	Very low risk	Low risk	Low risk	Low risk
		Very slight harm	Slight harm	Moderate harm	Great harm	Very great harm
	Harm potential of the information contained in the open data					

Give your overall estimate of the overall risk of the published data

- Very low risk = 0 points
- Low risk = 2 points
- Moderate risk = 5 points
- High risk = 8 points
- Very high risk = 10 points

The overall estimate is:

5. Costs related to data processing and analysis

With respect to the data being analysed, costs related to processing the data (possible anonymisation), analysing it, reporting it and maintaining it must also be identified.

The costs related to processing and analysing the data are mainly the personnel resources and working time used in the automation/anonymisation or analysis processes.

The costs related to reporting and maintaining the data are mainly the costs related to automation and post-publication maintenance (e.g. frequency of updating).

Assess and record the costs caused by publishing and maintaining the data

- Very high costs (more than X daily work units) [cities will define]
- High costs (X-X daily work units) [cities will define]
- Low costs (X-X daily work units) [cities will define]
- Very low costs (less than X daily work units) [cities will define]
- No costs
- The costs cannot be estimated

6. Overall estimate

6 Transfer your overall estimate of the benefits and risks from sections 2 and 3 here.

7 Use the table to make an overall estimate with respect to publishing the data.

The overall estimate with respect to benefits is:

The overall estimate with respect to risks is:

The estimated benefit of publishing open data	Very great benefit	Open	Open	Open, but consider limits on usage	Consider in more detail (concentrated support function)	Consider in more detail (concentrated support function)
	Great benefit	Open	Open, but consider limits on usage	Open, but consider limits on usage	Consider in more detail (concentrated support function)	Consider in more detail (concentrated support function)
	Moderate benefit	Open, but consider limits on usage	Open, but consider limits on usage	Consider in more detail (concentrated support function)	Consider in more detail (concentrated support function)	Do not open
	Slight benefit	Open, but consider limits on usage	Consider in more detail (concentrated support function)	Consider in more detail (concentrated support function)	Do not open	Do not open
	Very slight benefit	Consider in more detail (concentrated support function)	Consider in more detail (concentrated support function)	Do not open	Do not open	Do not open
			Very slight risk	Slight risk	Moderate risk	High risk
The estimated risk of publishing open data						

- Transfer your overall estimate of the costs from section 4 here.
- The assessment of the potential benefits and risks is the most central thing in the process of publishing open data.
- Especially in borderline cases (cf. the areas marked with yellow above), there is reason to consider the significance of the costs related to opening the data.
- The cost estimate can also be used to establish an order of priority for publishing different pieces of data.

The overall estimate with respect to costs is:

#MakeWithEspoo product family

The results of the experiments conducted in Espoo to create tools of development, management and consulting according to the City as a Service objective.

City as a Service

Frameworks

describe the background and theoretical framework of development

- Framework for the innovation management of ecosystems
- Framework for customer information knowledge management
- Reference architecture for municipal governance

Handbooks

provide models and examples for the use of city developers

- Handbook for open participation
- Handbook for co-creation
- Handbook for the production and utilisation of customer information
- Handbook for competence management
- Handbook for multi-channel public services
- Handbook for electronic customer service support
- Kuntakanvas

Application examples

present concrete implementations in Espoo

- Iso Omena Service Centre as an innovation platform
- Innovation showroom
- Data privacy and opening data securely
- KYKY - accelerated co-creation by schools and companies
- KIPINÄ - Connection map of operations and data systems



6 Aika



Euroopan unioni
Euroopan aluekehitysrahasto

Vipuvoimaa
EU:lta
2014–2020



Uudenmaan liitto
Nylands förbund

ISBN [number here] (print)
ISBN [number here] (PDF)