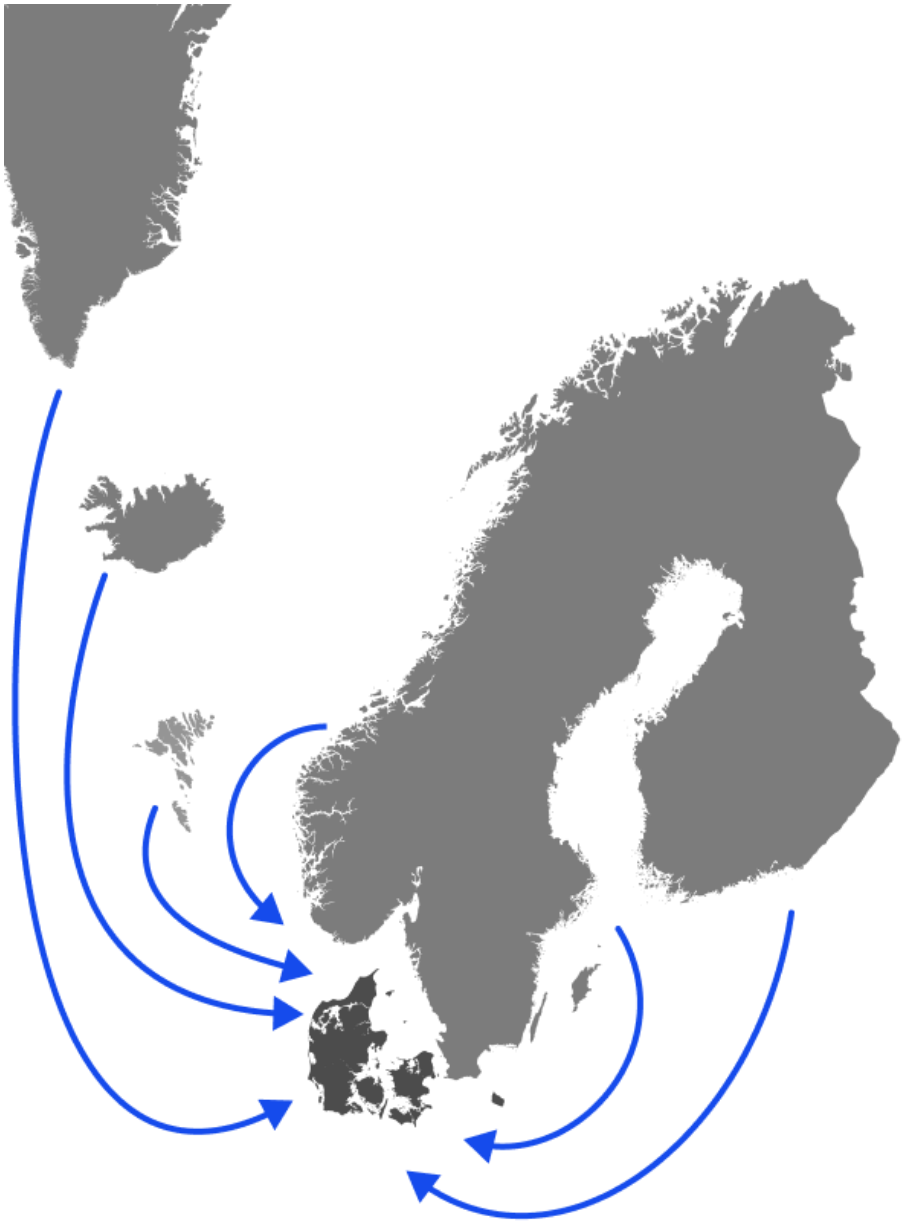
NILSTIERNA
ANALYS



Timely Nordic Welfare Indicators

The last of the first steps

Tom Nilstierna and Alexander Nilsson
Nilstierna Analys AB

2023-02-28

Table of Contents

A reading guide with a short summary

The present report has two principal parts. The first part, Main text, contains the main results of the project in an, hopefully, accessible form. For the more generally interested reader this should suffice. The second part, Technical text, contains text on technical issues that has been part of the project. The Main text touches occasionally upon the technical aspects that are handled in the Technical text. In the Main text there are referrals to the appropriate parts of the Technical text. The reader interested in technical details of the project is advised to read also the Technical text. The following short summary gives the results of the project in a concentrated form. The short summary would suffice for the reader that only wants/needs an overview of the results of the project.

- This project is the last step of the first steps of the creation of Nordic Welfare Indicators (NOVI).
- The project has mainly dealt with the need to use short period indicators in a timely manner.
- The project has investigated technical solutions for the handling of short period indicators.
- The main solutions are the use of automized processes for the handling of data.
- A prototype of the welfare monitoring system is published on the internet.
- Statistics harvested over the internet from the national statistical agencies has been used.
- One indicator, Earnings, has been created and tested in the prototype.
- Especially designated servers with web-capacities can and should be integrated.
- Further work in coming projects is necessary for a fully functioning monitoring system.
- An overarching cost-benefit analyses has been conducted that results in the tentative conclusion that the benefits probably are higher than the costs.

Main text

Motivation

The first question to answer is why we see the need to create a system of welfare indicators in general and especially Nordic welfare indicators. The answer might go somewhere along these lines: the new millennium has given us a few crises of magnitude. The financial crisis, the covid-19 crisis and the present inflation crisis are outstanding examples. We do not possess systems to closely follow and counteract such crises and, as a result, the policies that are established as answers to these crises are ill informed and not always to the point. Good information increases the chance of good policies being formulated.

It is a fact that the countries we live in produce high volumes of data on different aspects of the societies. The publication of statistics from the NOMESCO/NOSOSCO is included in this high volume of data. Data is not, in itself, information. There is need for more. The right data needs to be combined and presented in a way that aids, among others, policy makers to understand and, in best case, formulate good policies.

It is obvious that the political system(s) in the world have had problems counteracting the upcoming crises. There are probably several reasons for these problems. We believe that one reason is the lack of relevant quantitative information. Good policymaking needs accurate timely quantitative information. With a system for monitoring welfare the policymakers are equipped with aiding tools.

The choice to do this on a Nordic level is explained by the fact that it means a broadening of perspectives and makes it possible to consider the possible connections between societies.

With a full-fledged Nordic monitoring system of welfare, it can be that we will have a tool for debate and for choices on policies that in more precise ways can counteract an occurring crisis.

A developed and well-designed monitoring system may be a good aid for policy makers also in more normal times since welfare problems can be detected in groups of the population also during times of normality.

This report and the prototype

It is of some importance to note that this report, as obligatory as it is, is not the most important result of this project. A more important result of the project is that a prototype of the possible future monitoring system has been produced.

The prototype is a nearly fully functioning, although only minimal, version of the monitoring system we aim to build. Since it would be premature to have the ambition to within this project build the full-fledged monitoring system but still give an idea of the functionality of the ready-made product, we have created the prototype.

With the prototype we hope to start the important dialogue on what the publication of the indicators with interactivity can look like. Beside examples of possible indicators, we have in the prototype given different examples of presentations that can be chosen for the monitoring system. There are several different ways to present the indicators in visuals. Different groups of users have different needs when it comes to the presentation of facts in connection to welfare and changes in welfare. By starting up the prototype we hope to also start a dialogue of how best different information should be presented in visualisations.

The challenges

The creation of Nordic Welfare Indicators (NOVI) presents several challenges.

One challenge, of course, is timeliness, itself. In order to be of use as a fundament for policy choices the indicators need to be available not long after the events that created them. The availability depends on, at least, two parts; the first part being the measurement period and the second part being the administration period where the indicators are created from underlying data and published. It could even be stated that timeliness is a challenge constituted by the two challenges described above.

The fact that we want to create short period and timely indicators (with maybe monthly updates) presents us with another challenge. It would be impossible to work with the data the same way as the cooperation is currently doing it. A system that handles publication monthly needs to be based on programmed and automated procedures. The civil servants engaged in the cooperation still needs to set up the indicators, but they should not be involved in the actual daily running of the publication with updated data.

With NOVI we want to go beyond the mere publication of tables, often time-series. For the NOVI, we want to start a journey towards a contained monitoring system that gives the user a high degree of self-service (see *Power BI* on p.11) and with visualisations (see *Visualisations* on p. 11) that simplifies the interpretation of the presented indicators.

A decade of work

This report builds further on nearly one decade of work done within an effort to build a Nordic system of Welfare Indicators. The initiative from the Icelandic delegation to NOSOSCO to start researching the possibilities of such a system, must be considered to mark the beginning of this effort.

It must, though, be pointed out that this step is no more than one step of several steps ahead before the completion of a fully developed monitoring system of the welfare in the Nordic countries.

Two reports have been published during the years of work with the Nordic Welfare Indicators. One major report was published in 2016[1] and one minor report was published in 2022[2].

The report from 2016 proposes 33 indicators within 9 areas of welfare for a future welfare monitoring system. Mostly the statistics the report refers to is from the European statistical office Eurostat and is of yearly periodicity. In the report there is also mention of the need to move to shorter period statistics.

In the report from 2022 it was further emphasized that for a monitoring system to be useful it must consist of indicators that are of shorter periodicity – maybe monthly or quarterly. One other important conclusion was that the updating of indicators must be done in an automatic fashion.

The Prototype

The decision to build a prototype comes from the awareness that a fully working monitoring system has several dimensions to it. Beside the dimension of the actual indicators another important dimension is how the information gathered is presented. The prototype is built in Power BI® (see *Power BI* on p. 11) and consists of visualisations of central, although primitive, indicators of Nordic Welfare. The choice to use Power BI for the prototype should not be seen as a final choice of tools of visualisation for the monitoring system. Instead, the choice was made to make possible a speedy construction for early evaluation of how the visualisations should be when completed. The final choice of tool is to be made later in following projects in connection to Nordic Welfare indicators.

One purpose of building the prototype has been to examine how far the direct use could go of statistics found on databases that statistical offices provide. In the report from 2022 (the so-called Halfway report) from this project, it was stated that the direct use of such data is considered to be problematic since definitions and presentations of the statistics differs to such an extent that they are not easily compared. Despite that statement the project has used these statistics in order to better understand how far these statistics can be directly used and what, in a coming project, needs to be done for better comparability.

One note on comparability that has been discussed within the project is to differentiate between "level indicators" and "change indicators". Generally, it can be said that there are higher demands on data that are to be used for comparing levels than data that are to be used for comparing change. The indication of change is generally not as dependent of exact definitions of levels. This fact might make it easier to use statistics directly from the statistical offices.

Common Nordic statistical distribution system

Most Nordic Statistical Offices utilise – and further develops and maintains – a distribution system called PX (see *PC-AXIS (PX)* on p. 12) that provides services to access and work with data in a standardised manner with both a manual workflow with a file-based interface, and an automatable, programmatic interface – an interface we utilise in our prototype. This report may also refer to this system as *PxWeb*, the web interface to the dissemination system (see *PxWeb* on p. 12).

[1] A Nordic Welfare Indicator System (NOVI). Report for the Nordic Council of Ministers, 2016

[2] Nordic Welfare Indicators with timely data-The not to technical hallway report, Tom Nilstierna and the Task Force, 2022

Assignment databases

During the project it has become clear that it is possible to create assignment databases, that are especially designed to hold data in PxWeb format by assignment. This means that our own data can be stored and in connection to identification with a web domain can be reached through a web request. At this stage only a description what is possible can be made. We have decided to establish a PxWeb server in our own server environment. The Earnings indicator is distributed to the prototype through that server.

It is our belief that databases like this easily can be created in each country that takes part in this cooperation. We believe that we can only partly rely on data from the national statistical agencies. The cooperation needs to continue the work to create and publish indicators on its own. And for the handling of these data there is need for assigned PxWeb servers.

Scheduled updating

A well-functioning monitoring system with short period data needs a, more or less, automatic updating of the data it is based on. A handicraft updating of high volumes of monthly data is not a viable option. Within the project we have therefore experimented with a process that allows for scheduled updating of data. This essentially means that the updating demands little work by the organisation. The lion's part of the work will be done during the development of the system. See *Automating updates and handling errors* on page 10.

The indicators

It is not the purpose of this project to propose a full set of indicators for a future monitoring system. The purpose, instead, is to give central examples of possible indicators as such and how they can be part of a monitoring system with integrated tools for the end-user to control and choose for, for example, policy analyses.

The final set of indicators must be the result of a further "scientific" search and the work of the committees within the cooperation.

In the report from 2016 a set of indicators where proposed. The proposal contained yearly indicators. Since there has been no implementation of these indicators in an actual monitoring system, it was our original intention to, as far as possible, include them in the work of this project and to try to include them in a publication using web requests. For different reasons this effort has not been fruitful. There are indeed possibilities to harvest needed statistics at least from EUROSTAT. That would open for including these statistics in the prototype. On the other hand, we experience that the statistics available have in some cases changed since 2016, making exact copying impossible. This situation has made us decide to abstain from the effort of publishing the original yearly indicators. If there is a remaining need for yearly indicators, we propose that a new set of such are developed and published in combination with the short-period indicators that are the core of the monitoring system.

The truly investigative side of the prototype is the containment of the short period indicators that have been harvested from the national homepages. Not all indicators are fully covered in every country. Even so we have chosen to include some indicators of interest for our purpose in order to start a process for richer coverage.

During the planning of the project there have been calls for regional indicators. When possible and suitable indicators on a municipal level have been included.

The prototype contains short-period indicators covering the following areas:

- ✓ Labour market participation with unemployment.
- ✓ Earnings.
- ✓ Population variations.
- ✓ Transfer dependences.
- ✓ Daily deaths.

The experimental indicator "Earnings"

As an experiment and as a part of the prototype we have also chosen to create our own indicator on "earnings". The indicator is of course very "rough", and it must be developed for future use. It is our belief that a developed Earnings indicator will be central in a fully developed monitoring system. This belief stems from the fact that the earnings of the population probably are affected in nearly any form of crisis.

In most member countries the tax authorities receive monthly reports from employers. This data can be, and is, used for the creation of statistics. For the most part the data has been used for developing labour market statistics.

For our experimental Earnings indicator we have decided to ask for statistics that monthly gives earnings, that is incomes from employers to employees for actual work, that show incomes as close to renumeration for work as possible, in the population 16–65 for both sexes in their home municipality. This gives a regional aspect also. See *Experimental "Earnings" indicator* on page 10 for more information.

Portals with data

The Nordic statistical offices have solutions that makes it possible to easily harvest statistical figures with Application Programming Interfaces (APIs) (see *Data portals* on p. 11) and perform automated updates when new data is available. Nonetheless numerous interesting data is available in databases outside these containers. In the Nordic countries it is common that other state agencies are responsible for publishing statistics in the fields they are responsible for. The databases they are responsible for contain for us interesting data.

In principle it is possible to in the prototype include even more statistics from these portals. Since the aim of the prototype is to give a "glimpse" of how a monitoring system might look like we have chosen to restrict the number of indicators to a few and to elaborate a bit more on the presentation of them. In a later stage one option is to use these portals to investigate the possible extension of indicators. This work is probably best done with the involvement of committees within the cooperation.

Beyond this report

As mentioned now and then, the report that is now presented with the prototype is, hopefully, not the end of the journey with Nordic Welfare Indicators (NOVI). In the coming years, there is need for more work in the field to make NOVI a good tool for welfare monitoring.

In the, so called, half-way report from this project some possible directions for such work were proposed. Let us just, in the light of what has been here reported, stop by these proposals for a deeper understanding of what the monitoring system might be capable of in a few years' time.

Data on transfers to make possible the creation of gross disposable income are to be included

Earnings complemented with different kinds of transfers makes up the totality of individual incomes. The addition of these transfers makes it possible to construct indicators that show the composition of disposable income for different groups in the societies, thus making it possible to indicate the degree of dependence in different groups.

The inclusion of data on Subject Wellbeing (SWB) is to be considered

Beside on European level at the present only a few member countries do their own studies on SWB. First it is to be considered if and how these could be integrated into the monitoring system. Second it should be considered if and how a broadening to more countries could be realized.

The inclusion of health indicators

There are several ways to measure health. The indicators to be used are probably several and combinations of these are probably needed.

Use of developed hypothetical households for forecasting

The, so called, Life Situations within NOSOSCO are important tools for understanding the welfare systems in our countries. They quantify the change of disposable incomes when households go through changes of different kinds, like unemployment, sickness or childbirth. The same technique developed with quantities of numbers of households could be used to forecast both the supporting effect and budgetary effects of different policy options of rules in the tax-transfer system during a crisis.

The creation of micro databases for the variables that can give the sum of gross disposable income are to be built

Micro databases give a fundament for flexible aggregation and reporting. They create the possibility for analysis where groups can be formed in a flexible way. Using micro databases opens for the possibility to rearrange, and establish new, groups in the society depending on differing needs of analysis.

The use and development of microsimulation models (MSM)

There are MSMs in most member countries. These have been used in a common project within NOSOSCO. The traditional use is calculations of budgetary and distributional effects of changed rules within the tax benefit system. Such models could be used for distributional analysis and forecasting as a result of changes in tax benefit rules during a crisis. When policy analysis is performed MSMs can be used for analyzing the budgetary and distributional effects of the policy options under scrutiny.

Costs and benefits

Although this report cannot be the result of all the projects connected to NOVI, it is of value, at least in an overarching fashion, to ponder over the costs and benefits of the fully developed monitoring system we are about to construct.

The basic idea when estimating costs and benefits in general is to make possible a comparison between the two and making it possible to determine whether a project is worthwhile or not. The measurement of “worthwhile” would be a situation where the total costs are lower than the total benefits.

Costs

Starting with the costs it is probably fruitful to divide the reasoning into two parts: cost for investments and costs for operation. The foremost reason for this is to know the cost of an investment should be dispersed over a longer time and operations costs are directly resource consuming.

Investments

The investments can, in principle, be of three kinds: physical capital (like machinery), architecture and programming.

The expense for an investment differs from its costs due to the time frame of the use of the capital it creates. The expense for a new car is much higher than the yearly cost due to the perennial use of it. The capital cost is to be broken down into two parts: the financial costs and the cost of depreciation.

Physical capital

The needs for investments in physical capital are probably quite small. There could, of course, be a need for computer and communication-oriented machinery. It is our belief, though, that such machinery, for the most part, already is in place within the organisation and in the organisations of the member countries.

Architecture

It is different with needs for architecture. These investments contain meetings and decisions on the formation of indicators to be used in the system. It also contains software to be used within the system. We will use different parts of the PX system. We will use one, or even several, visualisation tools. We will, very likely, use tools for the publication of the reports within the monitoring system.

Of these most costs will fall on meetings and decisions. There are also costs involved in the publishing of the reports, but these costs are probably best viewed as operations costs since they mostly involve subscriptions.

Programming

There is need for the programming of software for building a structure for the monitoring system. These costs can, as an establishing effort, be quite substantial. It needs to be pointed out, though, that this primary programming would entail an initial investment that can serve for a number of years.

Operations

The, more or less, daily running of the system is, of course, of interest. It has been a conscious choice within the project to strive for solutions that keeps operation costs low. We have strived for the investment in procedures that automate the operations as far as possible. Even so, it is important to try to pin-point the needs for operating a system like ours.

Subscriptions

Although, the software that is needed for the system is mostly free of charge, the operation of it involves subscriptions to some extent. The exact amount of such cost depends on the volume of our publication. If we are to publish within a narrow circle of, say civil servants, the monthly costs can be as low as a few hundred Danish crowns each month. If we, on the other hand, want to make our reports available, in a dynamic format, for several hundred employees in the public sector in our member countries, we might face monthly subscription costs amounting to several thousand Danish crowns.

Supervision with error handling

When fully built, our monitoring system will probably include a few hundred tables that are to be monthly updated. The idea is that the updating is to take place through pre-set schedules and as such will, in principle, not need human handling. There is a high probability, though, that some updates will not work. We have experienced exactly that during the project. The updating errors can also be of different kinds (see *Automating updates and handling errors* on p. 10). One type of error

stems from the fact that the web requests done are faulty. That can be corrected over time. Another type of error is caused by changes at the delivery side. One example is when a table is withdrawn or changed somehow. During the project we experienced this kind of error especially when trying to include EUROSTAT tables proposed by the NOVI 2016 report.

Any kind of error during an update needs to be handled and it is close to certain that they will occasionally occur as a result of changes on the delivery side.

As an investment there will be programming covering error handling. It is unavoidable, though, that the actual handling of errors will include human work.

The error handling is a part of the supervision of the system and is probably needed on a daily basis. It is also probable that this supervision is to be part of an organisational function, like the present secretariate of NOMESCO/NOSOSCO.

Benefits

Already the costs of the system are difficult to calculate precisely. The benefits of it are even harder to calculate. In a sense the difficulties are similar to calculating the benefits of a military defence. In times of peace, the benefits are not obvious.

This, of course, does not mean that any monitoring system is beneficial, just as any one military defence is not beneficial.

There are a few dimensions of our proposal that are truly beneficial.

Timely data

The fact that we will be able to with as much as monthly updates report with vital indicators on the welfare situation in the Nordic countries is probably the most beneficial aspect of our proposal. Time is of essence especially in a crisis.

Collected information

Another benefit is that the information is collected and presented in a unified fashion. The fact that we now live in a world where data, like statistics, is abundant does not in itself mean that the information is at hand. The monitoring system brings together information that we believe can be of use when trying to understand the welfare situation in a certain time period. This of course does not guarantee that all information needed for that understanding will always be found there, but we do believe that important core information will be there for the analysis of many situations.

We will in one place have the relevant information on welfare from all Nordic countries. This helps broaden the perspective and helps in analysis on the interplay between national and international aspects of welfare developments.

Presentations with visualisations

Even the fact that we have a collection of data and information is not enough to make welfare situations comprehensible. Tables of data are not enough. We will be able to establish a monitoring system that with the aid of visuals of different kinds makes welfare situations easier to interpret and to start making policies in different situations.

Serendipity and side gains

The expression “serendipity” in the English literature points to situations in life when projects and human undertakings in general lead to gains, more or less, unintended. It is our belief that our project can spread benefits to other parts of the Nordic statistical cooperation.

The statistical work that is done within NOMESCO/NOSOSCO is probably the most obvious area to which benefits can spread. The actual handling of data within those committees are, in spite of the last few years of improvement, a cumbersome, and costly, process. At least some of the artefacts brought into the light of day by our project may well be of use for lubricating the further development of the statistical work within the committees.

Comparing costs and benefits

The moment of truth is of course when costs and benefits are to be compared. From a Nordic perspective, that is from a perspective of the totality of all affected in all the Nordic countries, there is value in what we are proposing only if the benefits are higher than the costs involved.

At this point in time, it is not possible to quantify costs and benefits. The question then is if we, in spite of that shortcoming, are able to say anything more of the aggregated costs and benefits in order to deliver a prediction of the balance. Maybe we can give some idea.

The costs are concentrated to capital costs in connection with investments in the system. The benefits are spread to users of the system but also to, at least, other parts of the Nordic statistical cooperation. It is therefore, as a tentative conclusion, highly probable that the benefits will be higher than the costs.

A more specific and quantitative calculation, at least when it comes to costs, can be presented at a later stage when the effort to build a Nordic Welfare Indicator System has matured further.

Technical text

Automated data collection and reporting

Automating the report system for NOVI can be – such as in our prototype – accomplished using digital data reporting services from sources such as the official statistical offices, and other government agencies with statistical reporting using web services.

Collecting these data into an environment for further reporting helps to minimize data processing labour efforts, although investment into the design and deployment of the automated collection and reporting system must be prioritized before such a system is viable in production-use.

Automation can be accomplished using web services such as PxWeb and other statistical data providers.

Experimental "Earnings" indicator

A Nordic Earnings indicator has very different degrees of maturity in the different member countries. In one country it was possible to simply download from a database. In another country we had to order a tailored delivery. For some countries we have for purely time-limit reasons decided to not pursue the actual gathering of data.

In order to do a proof of concept (POC) we have decided to establish a PxWeb server to simulate the possible stream of data for statistics on earnings that presently cannot be harvested through web-requests. For one country we have received the data in CSV format. We have thereafter with PxEdit made a table and put it on our own PxWeb server. As a last step we have from our visualisation system called upon the table with a web request.

Web services

Web requests to web services, or Application Programming Interfaces (APIs), such as PxWeb (see *PC-AXIS*) allow developers to work with data in their desired environments. Web requests are nowadays standardised with multiple workflows, such as Representational State Transfer (REST) that is built upon key principals defined by Roy Fielding[3] who laid out these core principals of architecting a software solution.

Utilising web requests in the context of statistical reports allows one to download statistical data from standardised web interfaces with the latest data on hand in an automated way, allowing for the development and publication of reports containing up-to-date information without the overhead of manual collection and processing.

Automating updates and handling errors

In the report prototype that is using web requests to collect statistical data, updates are largely automated, though all of these requests will need to be surveyed and maintained for continued working state of the updates. Managing changing or breaking URLs is a continued maintenance task that only must be performed upon the occurrence of errors that are automatically reported to the developer of the report system.

Errors are reported with an error code when an update fails, such as a missing or corrupted resource errors. However, no errors are reported in the case that no new data is found – in such cases, the data set is simply not updated. It is up to the developer how the data is updated from these requests,

[3] Doctoral dissertation *Architectural Styles and the Design of Network-based Software Architectures* by Roy Fielding, 2000.

such as rolling updates that removes the oldest observation and adding the latest, additive updates that adds new observations, or simply keeping the same time interval as originally requested.

It is our experience, though, that creating a well working process for scheduled updating is quite demanding. The web-requests that are the basis need to be well designed and the supplier of data might make changes that disrupts the automate. During the project it has therefore become clear to us that there is need for an error-handling mechanism within the process of scheduled updating. So, within the prototype such error-handling, although rudimentary, is also a part.

Data portals

Data portals are a collection portal of available data providers through web APIs. These portals provide services for accessing data through a range of data formats, frequency of updates, and organisations.

Through initiatives in the countries also these data are often reachable using data portals that have been created to simplify the capture of the data all over the societies. The portals contain the API: s needed and with a few programming steps it is possible to use the API: s and even automize the updating of data.

During our work we have come across and used the following data portals: *datavejviser.dk*, *data.norge.no*, and *dataportal.se*.

These portals guide the use of web-requests within a wide range of areas, not least the welfare area. In this work we have chosen to include some statistics that are collected with some of these portals as a starting point.

The data portals previously mentioned present data providers of official statistics in the countries Denmark, Norway, and Sweden; other data portals may provide non-official statistics providers – one must be mindful of what data providers are to be trusted as official.

Visualisations

Visualisation of data is as its name suggests: Reporting of complex data in a visually digestible format. There are many manners of and systems for visualising data, and the visualisations can range from non-interactive to fully interactive to the end-user. Visualisation of data helps to engage the audience with targeted reporting and customizing the data reported, regardless of how the data looks in the background.

A multitude of solutions exist to digitise the processing and reporting of information, and these solutions vary greatly in ease-of-use, learning curves for the developer and so on. Power BI by Microsoft enabled an intermediately skilled user to accomplish visualisations without the need for advanced programming of data collection, processing and finally the visualisation itself.

Programmatic solutions to visualisations in languages and frameworks require a more technically competent developer to work with the underlying data and to visualise them using the chosen visualisation solution – whether using a premade solution, or by developing an in-house solution.

In our prototype we utilise Power BI for our visualisation and self-service platform.

Power BI

Microsoft's Power BI (Business Intelligence) is a visualisation platform that allows for data collection from a range of sources for use in visualisations in reports that can then be delivered via web services

to end-users. Power BI reports allows for interactive elements that enable the end-user to filter and slice the data as desired for their use.

Our Power BI report prototype uses web requests to the PxWeb APIs provided by the statistical offices in the Nordic countries, thereby collecting Nordic statistics in one environment enabling us to create interactive indicators with the latest statistics, automatically updated when new data is reported.

Self-service

A key capability of the Power BI platform is the self-service functionality of the published reports – users access reports at their own volition in a user interface on a software-as-a-service (SaaS) platform with possibilities of authorization in order to protect access to reports to their key target groups.

Licensing agreements

An aspect that must be considered before the cooperation decides to deploy a reporting system based on the Power BI platform is the Microsoft license agreement models related to Power BI, user authorization systems such as Active Directory for self-service and access, and other related services.

Technical support service-level agreements (SLAs) must also be considered whether the cooperation technical support staff should have agreements with Microsoft for support if in-house staff cannot solve issues on their own.

PC-AXIS (PX)

PX is a system for disseminating statistical data through sub-systems such as PxWeb– these systems are used by most of the statistical offices in the Nordic countries. PX is built upon two types of data management systems – text files (flat-files) and relational databases. This flexibility allows for fast deployment and ease of maintenance based on the requirements and budget of the organisation that seeks to deploy the system.

PxWeb provides a web interface to the PX system, an interface that allows for accessing data programmatically through web requests – this is called the PxWeb API. Two forms of data access are available: payload request with table and variable definition in a JavaScript Notation Object (JSON) format, or by using PxWeb's *saved queries* wherein a query is specified in the PxWeb interface and then saved as a specific request URL, storing the query on the PxWeb server.

PxWeb

In most Nordic countries the so called PxWeb® is used not least by the National Statistical Agencies. This system makes it possible to access statistics in databases through the web with the use of so-called web requests. It is not necessary to use PxWeb for the acquisition of data through web-requests, but the nearly total coverage of PxWeb in the Nordic countries has made the use of it a pragmatic choice for us. In the few instances where such usage has not been possible the least diverging solution has been chosen.

PxEdit and PxJob

PxEdit – developed and maintained by Statistics Finland – allows for working with PX-files (a file with table definitions and data as part of the PX databases) in a graphical user interface. This tool allows for creating, editing, and merging, tables and table metadata.

PxJob is a command-line interface with similar functionality to PxEdit, with a focus on automatable tasks for tables and databases.

Assignment databases

The wide use of PxWeb among the Nordic National Statistical Offices has offered a platform for the creation of databases not already created by the statistical offices. For the further development of the Nordic Welfare Indicators, it is highly probable that there will be need for statistics not already created. There will be need for statistics particularly designed for the use within the Nordic Welfare Indicators system. For the prototype we have created a very primitive indicator showing monthly earnings from labour.

In an effort to learn more about these assignment databases within an isolated, project-specific environment, we established a PxWeb server in a development environment of our own. This endeavour helps to educate us on the requirements and considerations for deploying a database for the cooperation.

PxWeb databases can be deployed and maintained in two fashions:

1. Automating data aggregation from a selected portion of source tables from the statistical offices of the member countries.
2. Mass-importation of PX-file databases from all member countries, sorted into top-level country-based databases.

Maintaining a database on a Nordic-level PX system (PxWeb-servers, data management tools, etc.) will necessitate labour efforts, such as administration of Information Technology-systems (IT-systems), development of data collection and aggregation tools, and general upkeep of all the former. Designing this all-encompassing system requires a separate project with requirements and scope to be further discussed and laid out by the cooperation.

Data formats

Data formats are different file structures that allow for structuring data, as laid out in the subsections below. Understanding this section is not vital, however having an overall understanding of the data representations used in the prototype in this project, and perhaps further products in projects that follows, may aid in understanding the references to these formats.

Excel representation

Above two formats will, in Excel, look like this:

	A	B
1	COLUMNA	COLUMNB
2	VALUEA1	VALUEB1
3	VALUEA2	VALUEB2

As one can then hopefully understand by looking at the following two data representations that they are, in essence, the same as the Excel format; matrix-like data is possible in both JSON and Excel, with CSV/TSV being a flat text-file representation of Excel-files (with many of Excels functionality missing in the text format).

JavaScript Object Notation (JSON)

JSON-files are files structured in a key-value (KV) structure with allowing for nesting of KVs. A part of a file may look as such:

```
[
  { “COLUMNA”: “VALUEA1”, “COLUMNB”: “VALUEB1” },
  { “COLUMNA”: “VALUEA2”, “COLUMNB”: “VALUEB2” }
]
```

Comma/Tab-Separated Values (CSV/TSV)
CSV/TSV-files are text-based files with delimited (or separated; by commas, semi-colons, or tabs) data in a matrix-style column-row structures, and this may look like the following:

```
COLUMNA,COLUMNB
VALUEA1,VALUEB1
VALUEA2,VALUEB2
```

Programmatic frameworks for visualisations

Python is a general-purpose programming language commonly used in data science-applications such as data processing and visualisation.

R and SAS are programming languages and environments specifically designed for statistical data processing and visualisations.