The Basic Data Programme – A Danish Infrastructure Model for Public Data

Authors

Nicolas Lemcke Horst
Danish Agency for Digitisation
nho@digst.dk

Sara Bjerre
Danish Geodata Agency, sabje@gst.dk

Morten Lind
Ministry of Housing, Urban and Rural Affairs

Line Hvingel
Aalborg University, hvingel@land.aau.dk

In the autumn of 2012, the Danish government, Local Government Denmark and Danish Regions signed the agreement about 'Good basic data for everyone – a source of growth and efficiency'. The Basic Data Programme was born, and Denmark had an official and authoritative infrastructure model for spatial information. Work has run in parallel to the implementation of the INSPIRE Directive, and although the two development works have had each their trajectory, there has been, and still is, good synergy between the two infrastructure projects. This article presents the Basic Data Programme and takes stock of the implementation of the same.

Keywords: Spatial information, infrastructure, digital administration, INSPIRE, Master Data Management, eGovernment, Information Management

Introduction

Geographical basic data are localised information about, among other things, cadasters, buildings, road systems, watercourses and lakes. These so-called map data are generally of a high quality in Denmark, but they could be utilised even better in both the public and the private sector.

Climate adjustment is one of the areas where the data basis for efficient collaboration is being improved. By establishing unified, nationwide basic data for watercourses and an updated common elevation model, climate adjustment can be made more effective and cost-effective. By simulating water's flow paths on the surface, it is possible, for instance, to point out areas that are at risk of flooding in case of cloudburst.

Another example of how basic data can improve the current situation is house purchases. Easy and free access to data can give house buyers a more smooth case process, save time and money in the financial sector, and be of benefit to the entire society. When, for instance, a citizen buys a house, this triggers a flow of data between a number of players – estate agent, bank, mortgage company etc. As things stand today, the financial sector typically develops its own systems with data from a number of sources, which are not always up-to-date, in order to avoid paying for using public data. And they only download the data

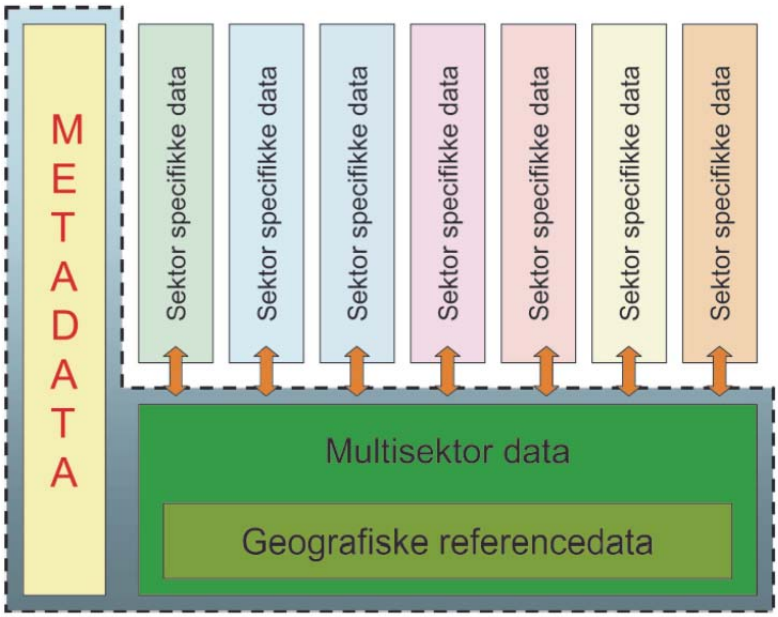

Figure 1. The Committee's preparation of a conceptual framework for basic data (basic data are contained in the box surrounded by a dotted line).

that are absolutely necessary. This means that administrative procedures are not optimised, and that time and money are spent developing systems that would be superfluous if data were flowing freely. Public authorities also spend a lot of resources on checking whether the private/public organisations are paying the required fees for the use of data.

The Basic Data Programme is described and implemented via 'Good basic data for everyone – a source of growth and efficiency', published by the Danish government and Local Government Denmark in 2012. This means that Denmark now has an infrastructure model that defines how basic data, reuse of data, and the connection between data in the public digital administration in Denmark look. Denmark has worked with digital administration for almost 40 years. For instance, the Danish Building and Housing Register (BHR) dates back to 1976, and in 1989, the cross-reference register was invented, making it possible to link and connect many different registers using selected keys. The use of spatial information systems accelerated up through the '90s.

In 2002, Hanne Brande-Lavridsen a.o. published an article that reported that Denmark had a de facto infrastructure model for spatial information, as lots of digital maps, digital registers, data models and ministerial initiatives existed, but that there was no official guidelines or models for how the infrastructure was to appear (Brande-Levinsen a.o., 2002).

In 2004, the Committee for Innovative Thinking about Basic Data published a report called *Basisdata – forståelsesramme og analysemodel til kategorisering af basisdata* (Basic Data – a conceptual framework and analysis model for categorisation of basic data) (Servicefællesskabet for Geodata, 2004). The Committee's work was based on an objective stating that *"Unambiguous connections must be guaranteed, and uncertainty and doubt about accuracy and data basis must be avoided. Geodata must therefore be localised on the basis of authorised and well-documented basic maps or other georeferences (e.g. address and building coordinates). Such basic data must be easily accessible"*. The Committee presented a model that bears many similarities to the conceptual models behind INSPIRE.

The members of the Committee have also been involved in the development work surrounding the INSPIRE Directive, which has sparked synergy between

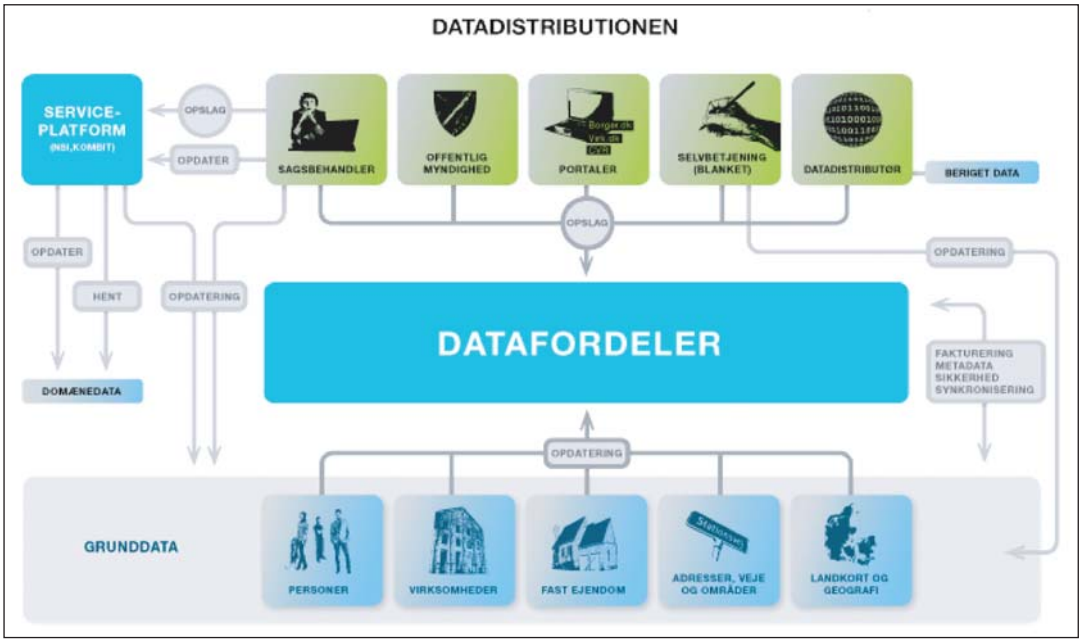

Figure 2. Establishing a common basic data infrastructure will ensure not only that data are made available for free and that they can easily be used in both the public and the private sector, but also that all basic data will be harmonised technically so that they can be used in digital case and business procedures.

the development works. The basic data report explicitly mentions that the recommendations are based on a view that corresponds with the thinking behind the INSPIRE Directive.

In 2012, *Gode grunddata til alle – en kilde til vækst og effektivisering* (Good basic data for everyone – a source of growth and efficiency) was published by the Danish government and Local Government Denmark as a practical outcome of the objectives contained in eGovernment Strategy 2011-2015 (the Danish government, 2011). The main idea is that the basic information that public authorities register about citizens, companies, property, buildings, addresses etc. is to be selected as basic data and used (reused) everywhere in the public sector, enabling the authorities to handle their tasks correctly and efficiently across units, administrations and sectors.

As the figure shows, the following basic data themes have been selected:

- Spatial data
- Address data
- Property data
- Company data
- Personal data

These themes cover more than 10 authoritative registers.

The most important targets for the development of basic data include:

- basic data needs to be as correct, complete and up-to-date as possible.
- all public authorities must use publicsector basic data.
- as far as possible, basic data (excluding sensitive personal data) must be made freely available to businesses as well as the public.
- basic data must be distributed efficiently, accommodating the needs of the users.

With the approval of the Basic Data Programme, all basic data are thereby, as a rule, to be placed at the free disposal of all public authorities, private companies and citizens. This means that basic data are a common digital resource that can be used freely for both commercial and non-commercial purposes. Free access to good basic data

Table 1. The seven sub-programmes and status of the same

Sub-programme 1	**Efficient property management and reuse of property data** By creating a silo-destroying and cross-cutting definition of 'particular property', property is to be handled in fewer registers in future, in a 'uniform and safe' way. *In sub-programme 1, data laundering of a series of data is in progress, and the solution architectures are expected to be concluded by mid 2014 so that necessary adaptations of the Cadastral Register, BDR and the Owners' List can be put out for tender. All agreed improvements are expected to be implemented in the course of 2016.*
Sub-programme 2	**Efficient reuse of basic data about addresses, administrative units and geographical names** Addresses, geographical names and administrative units (which serve as references for localisation) must and can be gathered and homogenised. It is a well-known fact that ambiguous names and addresses can be a costly affair in emergency situations and disasters, but in everyday life, too, one combined register of addresses and other things will be useful for the national economy on both a small and a large scale. *In sub-programme 2, address improvements are being carried out, the DAGI register* (DAGI = Danmarks Administrative Geografiske Inddeling = Register of the Administrative Geographical Division of Denmark) has been established, and the solution architecture for the Address Register is expected to be concluded by mid 2014 so that it can be put out for tender. All agreed improvements are expected to be implemented in the course of 2016.*
Sub-programme 3	**Unified basic data for water management and climate adjustment** Municipalities and the government (the Ministry of the Environment) will build the foundation for tomorrow's climate adjustment on just one data set about watercourses, which is based on the data set for the Unified Public Administrative Basis (UPAB), known under its Danish acronym, FOT. *In sub-programme 3, it is expected that an in-depth plan for the programme's completion will be ready by the autumn of 2014. The plan is to be confirmed by the program owners (the Danish Agency for Digitisation – the government, Local Government Denmark and Danish Regions).*
Sub-programme 4	**Free and efficient access to geographical data** As of January 2013, there has been free access for everyone, public and private users alike, so that geographical maps, cadastral maps, elevation data and more can be downloaded from the Danish Geodata Agency. *Sub-programme 4 has thus been completed in part; the part that deals with maintenance of the data, and which requires a more binding collaboration between the government and the municipalities, was approved by the FOT board in April 2014.*
Sub-programme 5	**Efficient basic registration of people and fewer duplicate registers** A thorough analysis of the basic registration of people has been launched in order to gain full clarity of possible solutions. This may lead to decisions about ensuring that each individual in Denmark is issued with an unambiguous personal identification key (in addition to the CPR number)**. Decisions may imply certain modernisation, e.g. of the duplicate registers that overlap to a certain degree, borrowing data from the CPR register. *It is expected that a political decision will be made about the modernisation of the personal data register before the end of 2014.*
Sub-programme 6	**Efficient reuse and sharing of basic data about companies** The registration of companies is to be expanded so that all companies, regardless of size, are issued with an unambiguous identification key; this also applies to foreign companies operating in Denmark. Data about production companies will be linked to geodata as a supplement to the authoritative address. *In sub-programme 6, great progress has been made with analyses and solution preparation: Most recently, the government has presented a growth plan for political negotiation, which may lead to the buying out of several sets of corporate data – accounting data.*
Sub-programme 7	**Common distribution solution for basic data – the data distributor** A unified data distributor is developed and established for distribution of all **basic** data covered by the programme. In the long term, the Data Distributor will also be able to deliver other data. The responsibility for the quality and maintenance of the data still rests with the individual registering authorities. *Sub-programme 7 is sticking to the schedule: One milestone was June 2013, when the tender material was forwarded to the five prequalified suppliers; the next milestone is when a contract is signed with the winning supplier. This will happen in June 2014.*

* Denmark's Administrative Geographical Division (The DAGV data set) is a standardised reference data set that describes and shows Denmark's administrative geographical division.

** The Danish Personal Identification number (CPR-number) is a national identification number, which is part of the personal information stored in the Civil Registration System (Danish: Det Centrale Personregister).

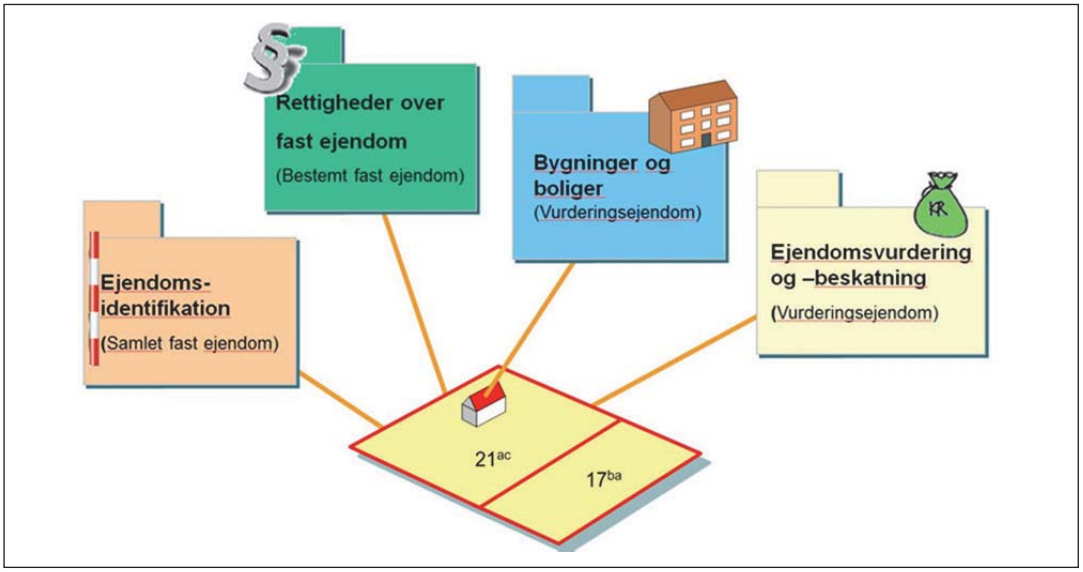

Figure 3. Today, three different property concepts are used in the Cadastral Register, the Land Register, the BDR and the Property Register, respectively. The completion of sub-programme 1 will see the term 'particular property' incorporated as one common term in all foundational registers.

for everyone is a good deal for the public sector and for society as a whole. Once the initiatives have been fully implemented in 2020, society's gains are expected to amount to 125 million EUR per year.

The Basic Programme consists of seven sub-programmes, which are described in outline below with a status indication for each of them.

Tidying up property concepts

Today, there is a very complex connection between property data in the primary property registers. At state-level, registration of property is handled by the Cadastral, Land Registry and Tax authorities. The municipalities also register information about property. The IT systems and procedures that are used in this connection are designed to solve specific tasks for the individual authorities. These systems and procedures were created at a time where dynamic data exchange between the property registers was not a given, and there are examples of the same information being registered and maintained in several places. As a consequence, the basic registers use each their own property concept and different keys for identification of property:

- The Cadastral Register uses the concept *Samlet fast ejendom (SFE)* (complete property), where the key is the SFE number. A complete property may consist of several plots of land, which are identified by the cadastral designation.
- The Land Register uses the concept *Bestemt fast ejendom* (particular property), where the key is a cadastral designation that may be supplemented by a serial number in relation to freehold flats.
- ESR/SVUR (MPR/NSAR = the Municipal Property Register and the National Sales and Assessment Register) use the concept *Vurderingsejendom* (assessment property), which has a 10-digit property number as its key.

Once the property data programme has been completed, processing and registering property in Denmark will be based on a cohesive and uniform basis, as the property area's basic data will be available in the authoritative registers – the Cadastral Register, the Building and

Dwelling Register (BDR), as well as in a new register of property owners, the Owner Register.

Via a unified public data distributor, all public and private users of property data will have access to relevant information about property in a harmonised, simple, up-to-date and cohesive way.

Information will be added to the Cadastral Register about freehold flats and buildings on leased land, and it will then contain information about all properties in Denmark, registered in a uniform way under one common property concept – Particular property. The Cadastral Register thereby becomes the authoritative basic register for all types of property. The Cadastral Register will be expanded to be able to register and display information about property that is currently under construction.

The Building and Dwelling Register (BDR) will remain an authoritative basic register for information about all buildings and dwellings, but it will include a reference to the Cadastral Register's registration of the property to which the building/dwelling belongs. The BDR will be expanded to be able to register and display information about buildings/dwellings that are currently under construction.

In future, the BDR will only contain information about buildings/dwelling, while address information will be transferred to a separate register (sub-programme 2 in the Basic Data Programme).

The Land Register will continue to contain information about the rights that are registered against the property. The registration of rights relating to properties that can be found in the Cadastral Register will be based on the information in the Cadastral Register.

An *Owners' List* will be established, which will carry over the Municipal Property Register's contents about actual property owners and any property administrators.

New municipal systems will be needed for calculation and levying of property tax, property contribution and rent/administration of municipal properties. The responsibility for setting up these systems, which will replace the property tax part of the MPR, rests with Local Government Denmark and KOMBIT (municipal IT community). The establishment of these systems, combined with the basic data registers displaying property data, is a prerequisite for shutting down the MPR as an inter-municipal main property register.

Effective addresses in public administration

Over the last 10-12 years, a number of public initiatives have launched a nationwide, geo-referenced data set that gathers Denmark's authoritative addresses. In parallel to this, a continual modernisation and simplification of legislation has taken place so that the regulatory tasks have been defined clearly. (Danish Enterprise and Construction Authority, 2012).

Today, the main challenge is that the rest of the public authorities are not using the authoritative address data sufficiently in their IT solutions and working processes. Many authorities use significant resources on the maintenance of their own address information, even though the best data are already available and maintained by the municipalities in the Building and Dwelling Register (BDR). Often, citizens and companies also waste time adding extra address information that the public authorities already have. The basic principle in the initiative is to establish a new, actual address register and make it mandatory for all authorities to (re)use these address data. At the same time, the quality of the existing address data is being improved, and data are becoming much more comprehensive, updated and easily accessible.

Each year, the municipalities designate approx. 20,000 new addresses and 500 new street names. When a municipal address authority is to register the new data, the task is spread across several different IT systems and registers. The improvements of address data and the access to them will benefit many users of addresses. Gone are the days of uncertain addresses, and the creation of addresses is gathered in one procedure. In future, the authoritative addresses from a new Address Register can be made available as a 'turnkey' component for all self-service solutions – public

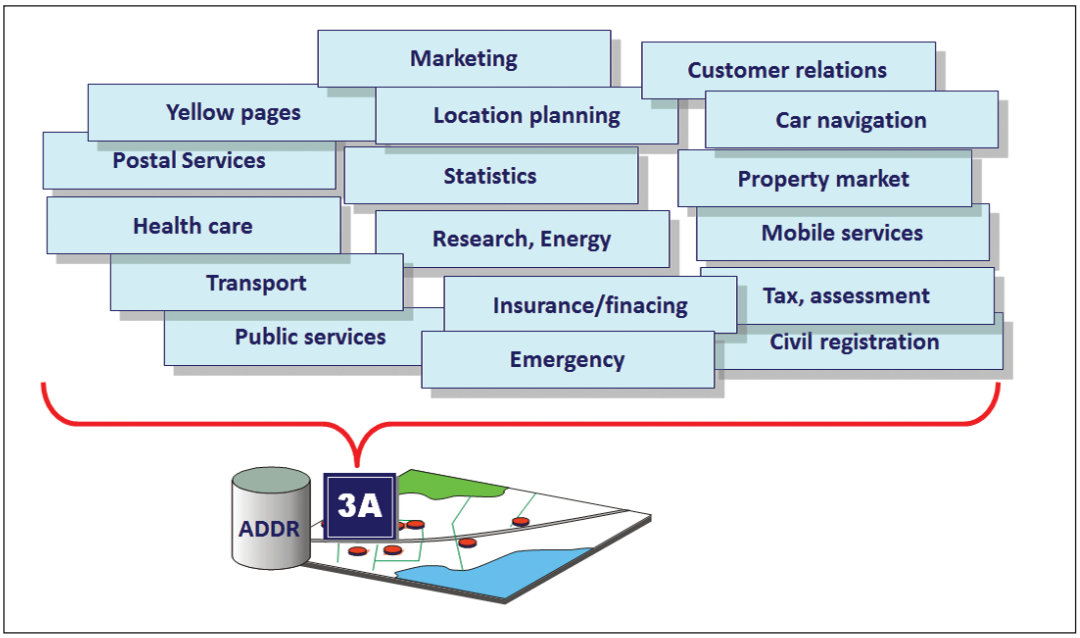

Figure 4. The vision of mandatory reuse of address data is that one authoritative registration of geographically coded street names and addresses will provide a secure key to thousands of uses in both the public and the private sector.

as well as private. The users will no longer have to enter the entire address, but will be able to select the address by means of 5-10 clicks. Often, the address can even be completed automatically based on CPR number (personal registration no.) or CVR number (company registration no.) – or using NemID (digital signature).

Common model rules are to make it easy to utilise basic data

Basic data are maintained by different authorities, but must be used together. As a data user working for a public authority or company, you will therefore need to be able to find quick answers to questions such as *What are basic data? Which information do basic data contain? And What relations are there between basic data?*

An interdisciplinary project has therefore been launched in the Basic Data Programme with the purpose of developing a common data model that will appear complete and cohesive to the data users. The common data model will be a sort of 'list of contents' of basic data. Common model rules are a crucial step along the way.

The model rules are to ensure both that basic data are modelled in a uniform way and are well-documented, and that they are of a quality and structure that make it possible to combine them across registers.

A common data model based on the model rules will make it easier for public and private data users and systems developers to gain an overview of basic data, adapt IT systems and phrase enquiries across basic data, thus bringing basic data into use at as low a cost as possible.

In the development of the model rules, it has been important to use internationally recognised standards, such as the modelling language UML (Unified Modelling Language) and ISO-standardised data types.

The model rules are, to a great extent, also based on principles and best practice from the EU collaboration on modelling of data that is covered by the INSPIRE Directive. Inspiration has been drawn from this for a number of rules for e.g. language, naming and documentation. Again, as in INSPIRE, one of the perspectives of the model rules is to make it possible for the data model to be translated automatically to data interfaces in for-

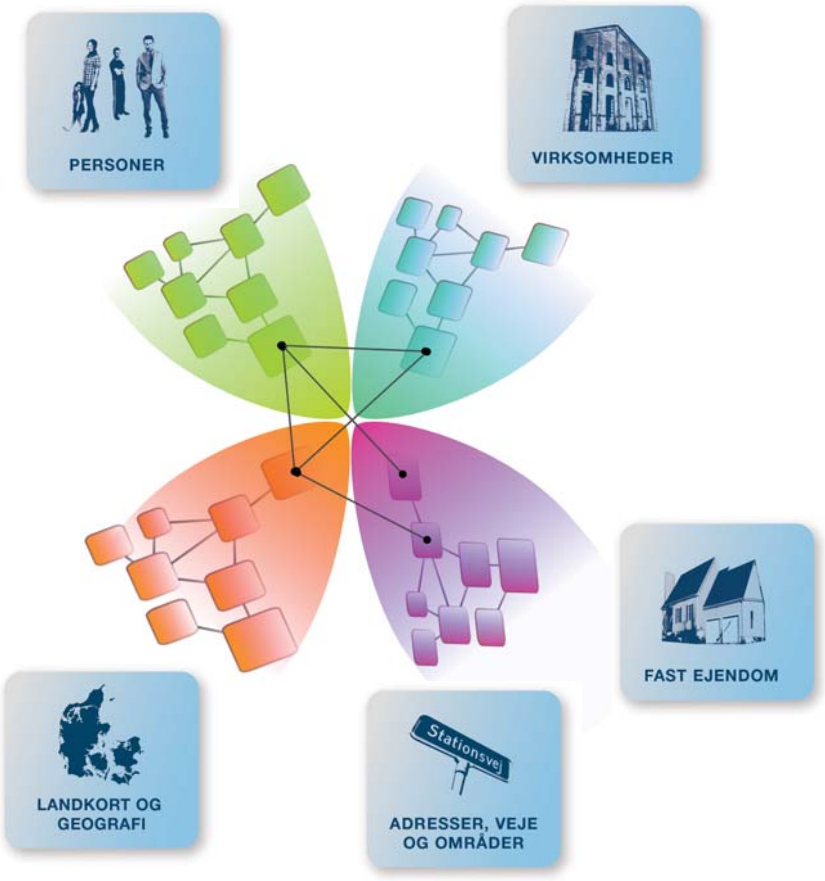

Figure 5. The basic data model provides a combined and cohesive model in a distributed administrative environment.

mats such as XML (Extensible Markup Language), JSON (JavaScript Object Notation) and GML (Geography Markup Language). The purpose here is to make the data model operational for systems developers.

The model rules therefore also pose demands about a number of general characteristics that all basic data should contain: *unique identification, bitemporality and status.*

Unique identification is to do with ensuring that all basic data can be identified uniquely according to a common, defined pattern, and that this identification must be unchangeable through the lifetime of the data. When basic data refer to other basic data, this must always happen by means of this unique identification. This will ensure cohesion across registers.

Bitemporality means that all basic data must contain information about when they were registered and for which period of time they are valid. In this way, it will always be possible to re-find information that public authorities had in their possession at a given time. The purpose is, among other things, to be able to document the concrete decision basis in connection with e.g. case handling.

Finally, all basic data must contain information about their administrative condition: For instance, are the data about a building that is under construction or demolition? Are the data proposals or approved data? The status information makes it easier for data users to sort through data and target enquiries.

In June 2013, the model rules were published on the Danish Agency for Digitisation's website with an open invitation to comment on the rules. Concurrent-

ly, a proof of concept was carried out where the rules were tested in connection with the development of a new data model for Administrative Divisions (part of sub-programme 2). Subsequently, the model rules have been adjusted and are now available as a version 1.0.

Over the coming months, the project will focus on the establishment of the display and the common data model and on maintenance of the model rules. More specifically, the basic data administrators will prepare data models for their basic data in correspondence with the model rules. The project coordinates the work and ensures that the data models are gathered and stored centrally. This will happen continually in connection with the implementation of the Basic Data Programme's sub-programmes. The display of the data model is to be made accessible in the way and in the formats that best support the requirements of the data users and systems developers. In terms of maintenance, it is essential that the model rules are to become a living document, and it must be ensured that the rules remain up-to-date in the future, in correspondence with requirements. The actual data model must also be adaptable in a way that guarantees that all changes are agreed and coordinated with and communicated to the right parties at the right time. The project is currently working on creating the framework and the organisation for this.

You can read the model rules and follow the common data model project here:

http://www.digst.dk/Loesninger-og-infrastruktur/Grunddata/Delprogrammerne/Faelles-datamodel

Concluding comments

The purpose of the Basic Data Programme is to secure an efficient use of public basic data through improved quality, free access and unified distribution of data. Better use of the public authorities' data is a means to significant streamlining in the public sector and to service improvements for citizens and companies.

By tidying up and improving public registers with key basic data and at the same time making these data freely accessible to all authorities, companies and citizens, it is possible to reap streamlining effects that will save about 35 million EUR per year in the public sector, due to fewer duplicated registrations, less parallel operation of registers and fewer administration expenses. At the same time, companies and citizens will be spared the trouble of providing the same information repeatedly to public authorities. Furthermore, free access to the top quality public basic data will save money and create new and better opportunities for growth and innovation in the private business sector to a value of almost 70 million EUR per year.

The vision behind the Basic Data Programme is to secure good, free and cohesive basic data that are updated in one place only and used by all.

REFERENCES

Brande-Lavridsen, H., and P. Daugbjerg (2002). *Omkring en dansk infrastruktur for stedsbestemt information (Danish infrastructure for localised information),* Geoforum Perspektiv, Volume 1. Geoforum Danmark. pp 5-10.

Danish Enterprise and Construction Authority (2012). *Improving an addressing infrastructure: the Danish experience,* in UPU The White Paper: Addressing the world – An address for everyone, pp 50-54. http://www.upu.int/fileadmin/documentsFiles/activities/addressingAssistance/whitePaperAddressingTheWorldEn.pdf

Ministry of Housing, Urban and Rural Affairs (2012). *Brug adresserne bedre (Use the addresses better).*

The Danish government / Local Government Denmark (October 2012). *Gode Grunddata til alle – en kilde til vækst og effektivisering (Good basic data for eveyone – a source of growth and effiency).*

The Danish government, Local Government Denmark and Danish Regions (2011). *The eGovernment Strategy 2011-2015.*

Servicefællesskabet for Geodata (2004). *Basic data – conceptual framework and analysis model for categorisation of basic data.*