

# **Addresses and Address Data Play a Key Role in Spatial Infrastructure**

**Morten Lind**  
**National Survey and Cadastre, Copenhagen, Denmark**

## **Abstract**

It is widely accepted that the address issue is of great importance in the world of Geographic Information (GI). Experience shows that the availability of full cover address data of a reasonable quality is often the key element that can “open up the world of GI”.

Using properly geo-referenced address data, every citizen, enterprise, institution or public authority will be able to "twist" their own address-related data from simple tabular data into multidimensional, spatial information. As thousands of private and public databases and registers carry references to address information, the potential of a simple data element that can release this "spatial superpower", is enormous.

A well-formed, public address system contributes to the physical infrastructure of a modern society, enabling ordinary people, postal services, rescue teams and utility companies etc. to navigate without coordinates. Proper address data can - likewise - contribute to the information infrastructure, pinpointing and labeling, in human language, all locations where people live, work, shop, entertain and educate themselves. Public address data is in many countries seen as a natural part of society's core data, on which a more sophisticated spatial infrastructure can be built.

## **Summary**

Section one of this paper describes the particular properties which characterize the address and the address system as a spatial reference system. The use of addresses as administrative keys in public administration in Denmark will be covered as a case story.

Section two will discuss the need for a consistent data model, when public authorities as well as private enterprises are managing address data, as well present some considerations and suggestions on how to model the address on a conceptual level.

Section three will illustrate different methods of using addresses as georeferences in a geographic information system (GIS).

Finally, section four will sum up some of the requirements which a modern society must meet in the management of addresses and address data to accumulate the greatest possible benefit.

## **1. Characteristics of the address as a reference system**

In the geographical world you distinguish between reference systems based on *coordinates* and reference systems based on *identifiers*. Addresses belong in the last category, which means that for instance the address “Rentemestervej 8, DK2400 Copenhagen” identifies a particular (more or less well-defined) location without use of geographical coordinates. (Figure 1-1)



**Figure 1-1:** The address system is based on identifiers, not coordinates

The identifiers of the address system thus consist of "names", that is, country, region, town, district, road and house or door number. It is a characteristic of the address system that it is structured hierarchically with a network of named roads as its backbone.

Place names make up a corresponding name-based reference system, as each place name indicates a particular locality or a particular area (town, part of town, village, lake, wood, district etc.) without using coordinates. A third example is the Cadastre where the name of the region (parish etc.) and the title number identify each plot in the area covered by the Cadastre.

### **1.1 The particular properties of the address**

The address system, however, is distinguished by a series of special properties from other name-based reference systems:

- *It is well-known and widespread:* Everybody knows the system, and addresses are without comparison the most widely used method of localization when storing information of the location of this or that person or business.
- *It is practical and logical:* The system is based on the way we get about: The hierarchical structure with town name, road name and house numbers, divided into equal and unequal numbers in rising order, is easy to find your bearings in on foot as well as in a car or on a bike or when looking at a map.
- *It is suitably detailed:* In the built up areas where we move around daily, the address system is so fine-meshed that we can find our way to the right front door or stairway entrance, solely with the aid of a road name and a house number.
- *It is visible:* Signs with road names and house numbers mean that we can read and find our way around the address system, and find that particular front door we are looking for. The combination of logical structure and visibility is perhaps *the* most special and valuable property of the address system.

It is these special properties which explain why the address is such a central component in GIS products and services: When we use addresses as a means of communication, we speak a "language" which people understand and can relate to!

We can assume that the address constitutes a potential geographical key in thousands of registers

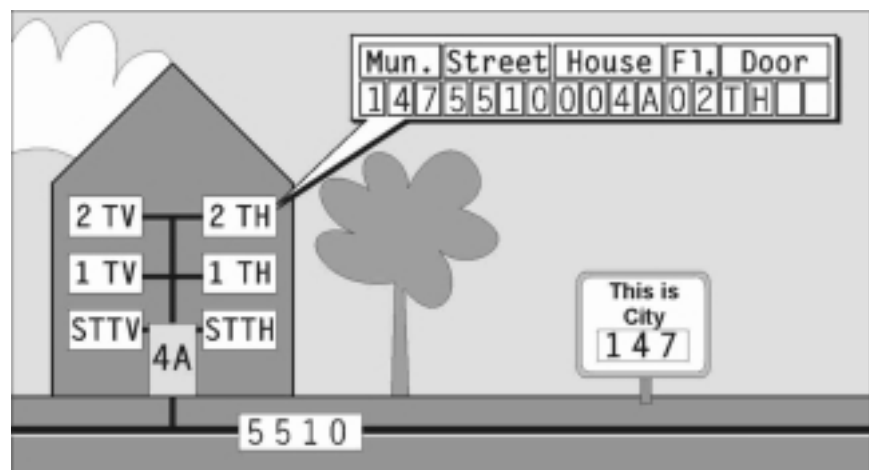
and databases within the private as well as the public sector, and that we will be able to link records in different databases with the aid of the address key to transform billions of text-based records into spatial information.

### ***1.2 A case: The address system in Denmark***

The address systems in Denmark (as well as in the other Nordic countries) are regulated by the legislation pertaining to the central civil registration and by rules and regulations in connection with property registration, e.g. registration of buildings, dwellings and property.

The Danish Central Personal Register (CPR) law from the late 1960s and the Building and Housing Register (BBR) law from 1977 form the basis of the public authorities' administration of road names and house numbers. The law governing the Building and Housing Register ensures that each individual dwelling is assigned a unique address designation which is registered by the municipality. At the same time, the law governing the Central Personal Register ensures that every citizen with a permanent dwelling is registered with his place of residence at a certain address. The demand that dwelling addresses must be unique entails that buildings with more than one stairway are addressed with a house number for each individual stairway.

Correspondingly, the insistence on uniqueness of reference entails that dwellings which are located on a stairway with more than one dwelling, must be assigned information about floor number and door designation. A Danish address designation is thus composed of the following elements: municipality number (3 digits), road code (4 digits), house number (3 digits + perhaps 1 letter) and floor (2 characters) and door designation (4 characters). This 20 year old address system is uniform over the whole country and includes towns as well as rural areas in all municipalities. (Figure 1-2)



***Figure 1-2: The "architecture" of the danish address format***

In the public sector it is a great advantage to use the official address format, because it reflects the administrative division of the country (by virtue of the municipal code). Moreover, the address format is based on digit codes and not on road or place names, where uncertainties about spelling, abbreviations etc. are often encountered. Finally, the format means that each individual element in the address is stored in a separate field which entails that address designations are easier to

handle and correlate computerwise. The drawback is, of course, that an official address designation will only be readable and understandable, if it is supplied with a municipality name and road name.

### 1.3 Benefits of using the address as an administrative key

There is no doubt that the function of the address as *administrative key* has already saved the public sector in Denmark considerable amounts of money. The existence of a reliable administrative key between the personal data area and the area of building and property data means that a connection can be created between various administrative public systems which handle amounts running into billions in the form of personal taxes, property taxes, health insurance, wages, pensions, benefit payments etc.

To cite an example, there has not been a traditional manual census in Denmark since the 1970s. On the other hand, a machine-based census and dwelling count is made several times a year based on the personal number and address systems. Corresponding profits are gained by private industry, the financial sector, for instance, which also uses the Danish "address key" to a large extent.

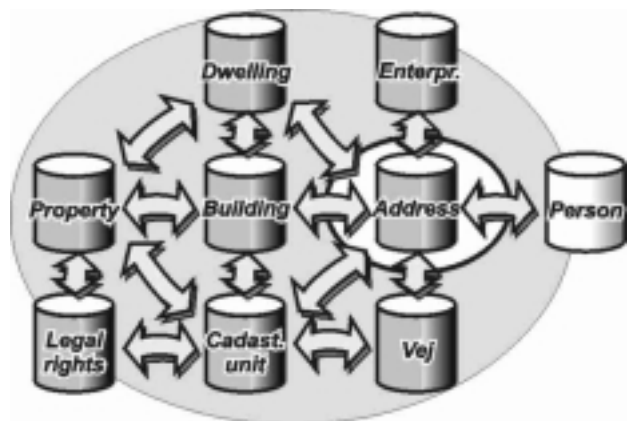
Concurrent with the increasing use of the address as *geographical key* by the public administration, the performance of a series of important daily public tasks, as for instance the fixing of school districts, the registration of patients and waiting lists, the holding of general and municipal elections, the performance of police duties, turn-outs for emergency calls etc., will grow appreciably more effective.

## 2. Datamodelling the address

In accordance with the above, Denmark ought to have the best possible starting point for using the address as an official, administrative "key" to link information from the various public and private databases and digital maps. Experience shows, however, that the existence of a common data format gives no guarantee of a sufficient consistency between the data contained in the various registers. (Figure 2-1)

As an example, practical tests have exposed the fact that address information in the National Business Register (CER), which is based on reports from individual businesses, does not correspond very well with the addresses found in the Building and Housing Register which is maintained by the municipalities.

**Figure 2-1:** Modelling the address at an independent entity or object type.



### ***2.1 Attribute or entity?***

The problem can be elucidated by viewing it as a datamodel problem on a conceptual level and posing the following questions:

- Should the address be regarded as an attribute, that is, as a characteristic of each of the entities in question: the person, the property, the building, the business etc. or
- Should the address be perceived as an entity, that is, as an independent object “related” to the individual person, property, building, dwelling etc.?

In the private as well as in the public sector, the predominant tendency is for computer systems to favour the first concept: With very few exceptions, address information functions as a simple attribute, that is, as a characteristic which is indicated in the database on a par with other “soft” characteristics like names, areas, conversion figures etc. without any kind of validation.

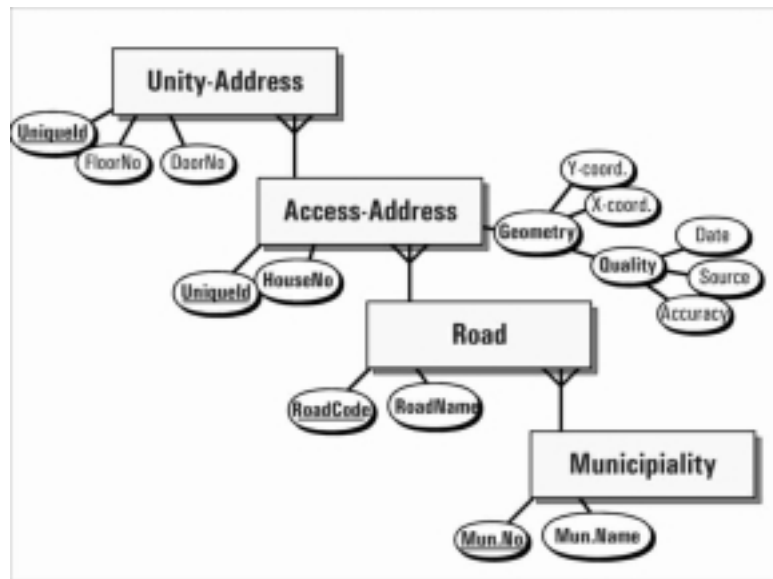
The drawback to this “attribute model” is that it prompts each individual computer system to build up its “own” address registration which has no connection with those of “other” systems. The result is that inconsistencies arise between the individual address collections, so that property, dwellings, businesses etc. which ought to carry the same address throughout the different registers, in fact do not. Furthermore, it has proved that the attribute model promotes situations where each administrative system creates its own frame of reference for the address concept, so different address sets are created, and divergencies arise in the context of the address.

### ***2.2 New datamodel: the address as entity***

In connection with a Danish project for the modernization of the Building and Dwelling Register, the idea arose of regarding the address as an independent entity, that is, as a separate object of registration to be treated with the same care as other comparable administrative designations, like for instance personal numbers and cadastral numbers etc.

The benefits of this new datamodel are that it strengthens the very concept of the address and positions the address more centrally as a common key object. The consequences will be, among other things, the following:

- The stock of addresses in a given society can and should be administered as a single quantity giving greater clarity as to which datasets should be used as validation basis in relation to other private and public registers and databases.
- The authority to handle addresses can be placed unequivocally with a certain authority or institution (in Denmark the municipal council) to prevent unnecessary duplication and to ensure that the exercise of authority is handled according to a well-defined set of rules.
- The address will, when it is established, exist “in itself”, and is thus no longer just an “appendage” to for instance a building or a plot. That means, among other things, that new addresses can be established and registered when practical, for example in a planning phase before a plot is parcelled out, or a building erected.
- The address can be registered and equipped with its own relevant data characteristics (for instance site coordinates, date of origin etc.) which can be reused in all data collections containing address information. (Figure 2-2)



**Figure 2-2:** The Danish Address system ER-model

### 2.3 A public address register

During the last couple of years, the government and the municipalities in Denmark have cooperated on a harmonization and quality control of the addresses in the public registers and maps. The purpose is to reach that set of addresses which will compose the “backbone” in an independent public address register. At the same time, work goes forward on establishing direct georeferences to the addresses in the register, using a set of geographical coordinates to each individual address appearance.

The overall objective is for the official address register of Denmark to represent a common resource in the geographical (geospatial) information infrastructure to the benefit of both the private and the public sector.

### 3. Georeferences: the address in a GIS

Before address data can be used in a geographical information system (GIS), the address designation must somehow be “translated” into a direct spatial reference with the aid of coordinates. The address "Rentemestervej 8, DK2400 Copenhagen NV, Denmark" which refers to a particular front door in a building in Copenhagen, could for instance be translated into the coordinates of a point = (6.178.913m; 777.142m) which indicates the approximate position of the front door in the building, measured in the international coordinate system UTM WGS84 (Zone 32).

When a website shows a blue dot on a city map as the answer to a query about the location of Rentemestervej 8 in Copenhagen, Denmark, that is the technique which is used. The website’s database system searches for the address in the address database, finds the corresponding coordinates, and calculates the position of the dot on the map.

The use of addresses in a GIS depends, however, on *which method* is used when the addresses are represented geographically in the system. In principle, this geographical representation can be structured in three ways: the polygon-based method, the network-based method and the point-based method.

### ***3.1 The polygon-based geographical representation***

The method implies that a polygon in the digital map represents a real world district with a corresponding set of addresses, for instance postal districts, parishes or constituencies. The cardinality between address and polygon is thus N:1.

This method is attractive in circumstances where knowing the exact locality of each address appearance is not crucial. The method is thus particularly apt for a presentation which aims at general information and analysis. If the chosen division into districts also carry other sorts of data, for instance statistic information about population etc., an analysis of the distribution of one's own data in relation to this information will be facilitated. Moreover, the method implies that only access to fairly simple (and thus inexpensive) geographical datasets is required.

### ***3.2 Network-based geographical representation***

The method implies that a digital network representing the real world road system is linked with information about which addresses belong to each line segment or perhaps each node in the network. The cardinality between address and road segment is here N:1 as well. In relation to the polygon-based method, the network-based method implies, however, that most GIS programs can arrive at an approximate locality of individual addresses by interpolation.

If the road network is structured topologically correctly, so that it reflects the possibilities of motion in the actual road system, the method has, of course, its great strength in applications where traffic on the road system is included as a parameter, that is, in route planning, traffic optimization, "fleet navigation" etc. As a general method for localizing individual addresses, the method implies an uncertainty which must be estimated for each individual purpose.

### ***3.3 Point-based geographical representation***

The method implies that each individual address is linked to a point in the map which represents the locality of that real world object which the address identifies. In Denmark, an address point will be placed with reasonable certainty in the right building, and as a rule within a distance of less than a few meters from the actual position of the entrance door. The Ordnance Survey's dataset "Address Point" for the UK operates with a similar accuracy. (Figure 3-1)

In contrast to the two other methods, the cardinality between address and point is typically 1:1. That means that in the point-based method, it is essential to know the completeness of the point theme in relation to the address set. In a dataset of good quality, one can locate each individual address and at the same time be certain that each address on the map corresponds to a real world object of which information can be sought in other databases.

The point-based method is, among other things, a prerequisite in circumstances where an address must be located exactly in a large scale map, aerial photo etc. (for instance by the police or the fire and rescue services), where it is crucial to know, if one or more particular addresses lie within or without a certain bounded area (for instance by neighbour hearings, pollution, contamination etc.). If the address points are combined with a topological road network, GIS solutions of very high functionality can be produced.



*Figure 3-1: Address points: Every black dot represents an specific "doorstep-address"*

### **3.4 Conclusion on address georeferencing**

The geographical representation of addresses depends primarily on which functionality and end result are required. Here as well, the means depend on the aims. Experience from the international geodata market shows that a broad range of datasets will be needed which can be used by private businesses as well as by public institutions to render their address-based data ready for GIS.

In the Scandinavian countries, the public sector has the possibility of further developing the address system, so that a transverse public address-based dataset of high quality can be established. Similar to the US Census' TIGER files, the various private businesses will be able to further develop and improve the quality of these public databases into new datasets directed towards different fields of application, market segments and user groups.

## **4. Managing the address reference system: requirements and challenges**

Any kind of reference system poses demands: certain common standards and definitions are called for. A responsible party must be nominated to exercise the necessary custodianship of the system. A legal framework must be established which ensures the best possible diffusion of the system, including formal agreements between parties and perhaps also legislation. The reference system must be well-documented, and the data and the parameters which describe it must be as freely accessible as possible.

On the basis of, among other things, the Danish and Scandinavian experience, these general requirements can be transferred to a public address system which will function as a common resource in a public information infrastructure.

### **4.1 Standards**

Experience in Denmark has shown that to regulate a data format for address identification by law is not sufficient, if at the same time free range is given to the individual address-carrying systems

to define addresses at a conceptual level. It is necessary, then, to find a common definition of the address system, of the address concept and of the address as an atomic minimum entity.

Newer examples of formal public standards for addresses are found in Sweden and Great Britain, whereas in Denmark, the address concept will be formalized directly in a new independent legislation on addresses in the course of this year.

At a European level, Working Group 3 of CEN/TC 331 "Postal Services" is working on the development of a five-part standard dealing with the topic of European postal addresses, their representation on postal items and their electronic exchange.

#### **4.2 Management**

In the Scandinavian countries, various legislation and agreements determine the local authorities (municipalities, city councils) as address authorities, e.g. as competent to assign road names, house numbers and other address identifications. In other countries, the postal service functions formally or in practice as address authority, responsible among other things for establishing house numbers and keeping them visible.

If the potential of the address key as a common information resource is to be utilized, it is a prerequisite that the conditions of authority are well-defined. It has to be quite clear who has the competence and the duty to establish, maintain and if necessary abolish individual addresses and other components of the address system.

In agreements or proper legislation it can be made a requirement that the central public registers for, for instance, legal properties, social security, insurance etc. or certain private services, as for instance the rescue services or the telephone companies' phone directories, must use the public address system.

#### **4.3 Access**

As the address system, as opposed to the coordinate-based reference systems, consists of a finite number of "discrete" components, a rational utilization of the system presupposes the existence of an inventory which describes the total contents of the address system, an address register or an address gazetteer. The Royal Mail's "Postal Address File" (PAF) is an example of such an inventory.

It is obvious that only the freest possible access to such an address file can ensure that the address system is diffused and used where it may benefit, and that no incompatible or alternative address datasets or databases are allowed to emerge. At the same time, it is important that private businesses and producers are allowed to base their activities or product development, (for instance within GIS, mapping, geomarketing, routeplanning) on data (including georeferences) from the public address system.

#### **4.4 Conclusion**

In spite of the many differences that the address systems in various countries and regions show, they still fulfill the same basic needs: To identify certain locations, where we live, work and educate ourselves, in a language which mirrors the way we get around: our common road network.

New points of view result in new challenges. The new angle on the address places it as one of the most central elements in tomorrow's infrastructure of core data and reference systems. New demands pose a challenge, especially to the public sector's co-ordination and administration of addresses.

### **References**

British Standard, *BS7666, Part 3: Specification for Addresses*, London, 1995, UK.

Cushine, J: "A British Standard is Published", *Mapping Awareness*, June 1994, p.40-43.

Lind, M., Lund Christensen, T., "The Address - a key to GIS", *Proceedings for Joint European Conference on Geographic Information*, March 1996, Barcelona, Proceedings Vol. 2, p. 868-877.

Raper, J. Rhind, D. and Shepherd, J.: *Postcodes: The new geography*, Longman Group UK Ltd., 1992, Harlow, England.

Swedish Standards Institution, SS 63 70 03, *Addresses*, 1<sup>st</sup>. edition, Swedish Standards Institution, 1998, Stockholm, (Publication only in Swedish: "Belägenhedsadresser")