

## **Statistics Norway's geocoding quality declaration for the Business Register**

Statistics Norway (SN) receives a copy of the Official Business Register every night. A challenge to use this in GIS analyses is the lack of X- and Y-coordinates. As the conceptual model shows, there is an Address Control B for the Business register against the official Cadastre, which holds the official addresses, buildings and properties in Norway. Unfortunately, this control is only a check of right spelling of street names and does not return a unique numeric address code for the exact address. Due to this, SN must do an in-house geocoding process to obtain coordinates for each establishment.

The Address Control A is returning the unique numeric address – see Address terms below. The population of establishments geocoded must be located in Norway.

### **The registers used in the geocoding processes**

The Business Register (BR) in SN is a copy of the Official Business Register mentioned above, containing enterprises and establishments. The aim of the geocoding processes is to attach X- and Y-coordinates to each Establishment (Local Kind of Activity Unit – LKAU) in the register.

The registers used for finding coordinates are all copies of Norwegian official administrative registers.

#### The Address Register

The copy of the official Cadastre holds both addresses, buildings and properties. For short, this copy is referred to as the Address register in this document. The basic statistical unit is also a part of this register. This register is used in the joining processes **A1, A2, A3, A4, A5, L1<sub>b</sub>, M1, N1, V1, W1, W2, W3, W4** and **W5**.

#### The Population Register

This is a register of all citizens in Norway. This register is used in the joining processes **R1** and **R2**.

#### The Norwegian Farm Register

The register includes all land and forest properties in Norway. Statistics Norway receives a copy of this register twice a year. This register is used in the joining process **L1<sub>a</sub>**.

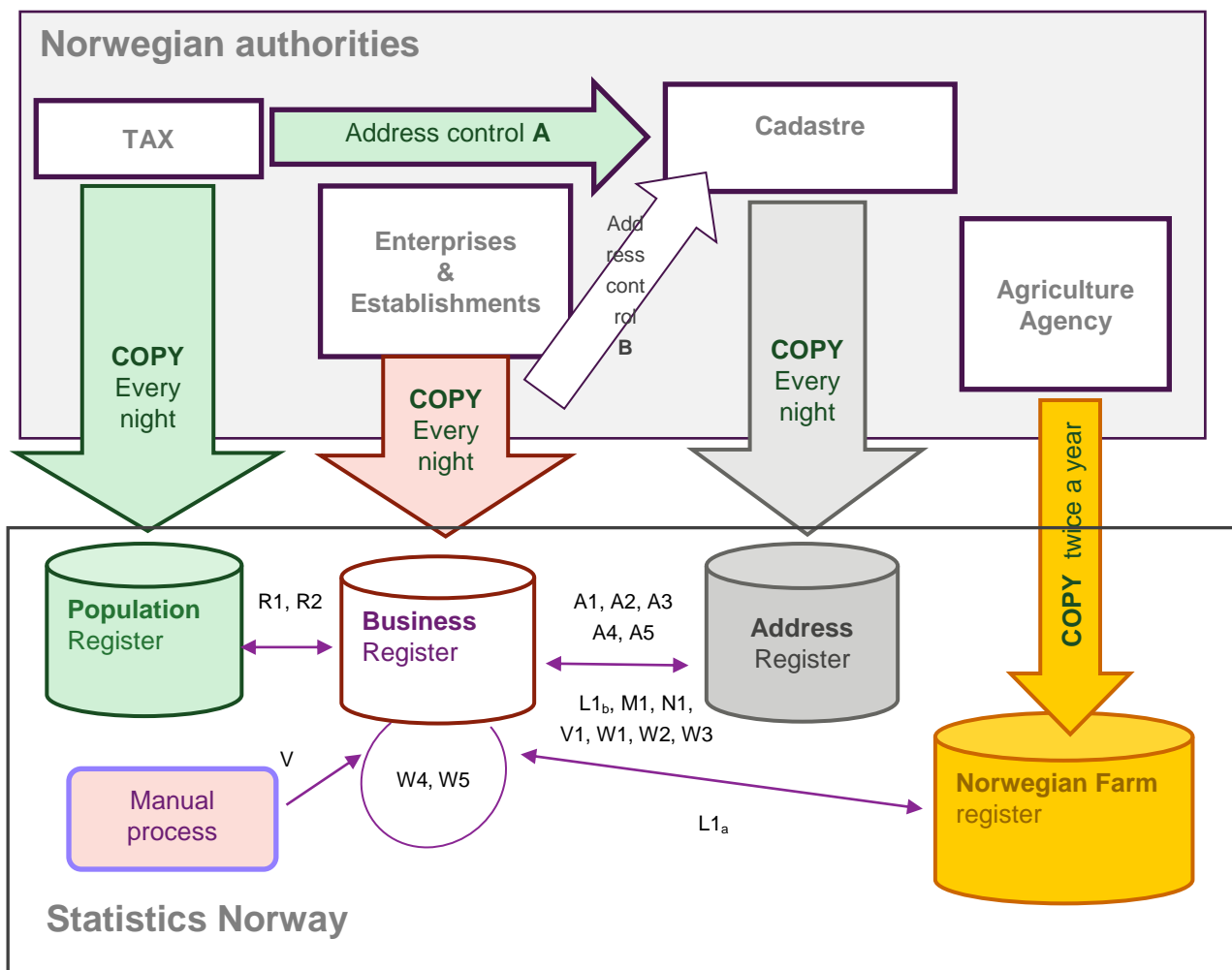


Figure 1. Conceptual model for geocoding the Business Register in Statistics Norway

### Address terms

Norway has a coherent and centrally maintained address system that consists of two different address objects; street address and cadastral address. They correspond to unique numeric addresses as shown in the table below.

Part of the street addresses	Characters in the unique numeric address	Part of the cadastral addresses
Municipality number	1 - 4	Municipality number
Road number	5 - 9	Cadastral unit number
House number	10 - 13	Property unit number
Letter	14 - 17	Leasehold number
Condominium number (0000 for street addresses)	18 - 21	Condominium number
Apartment number	22 - 26	Apartment number

Depending of the need for accuracy, the number of characters used when joining registers may differ. When the purpose is to geocode exact X- and Y-coordinates, the first 17 characters in the unique address are used. Basic statistical units are not part of the unique address.

*Modified address* is a term used in this document. This is street and road names spelled in a uniform way based on the original addresses attempting to adapt to street and road names in the Address register. Example is “Skolevegen” and “Skoleveien” (School road) where “vegen” and “veien” both are correct ways of spelling the word for *road* in Norwegian. Other ways of modifying for example street names are reduction of blank characters, hyphens and periods. Expansion of abbreviations to the

relevant word is also a way of modifying street names. Example of the latter could be “gt.” expanded to “gate” which is the Norwegian word for *street*.

### The joining processes

The joining processes runs every month, and the processes that provide coordinates are A1, A2, A3, A4, A5, L1, M1, N1, R1, R2, V1 and V2. These processes return also *basic statistical units* in most cases. The rest of the joining processes provide basic statistical units. The joining codes are given in a combined alphabetical and numerical order. I.e. the joining code A1 is likely the most correct address and the code W5 is the most uncertain address. Joining code XX means the establishment did not match at all.

Results from the joining in October 2017:

Joining process	Number of joins	Per cent
A1	491,978	80.97
A2	787	0.13
A3	1,508	0.25
A4	9,465	1.56
A5	424	0.07
L1	19,440	3.20
M1	6	0.00
N1	1,186	0.20
R1	22,876	3.76
R2	2,572	0.42
V1	2,049	0.34
V2	4	0.00
W1	132	0.02
W2	7,165	1.18
W3	910	0.15
W4	3,271	0.54
W5	43,079	7.09
XX	772	0.13
<b>Total</b>	<b>607,624</b>	<b>100</b>

Results from the joining in June 2018:

Joining process	Number of joins	Per cent
A1	499,581	82.51
A2	578	0.10
A3	821	0.14
A4	9,260	1.53
A5	272	0.04
L1	17,937	2.96
M1	6	0.00
N1	1,115	0.18
R1	20,467	3.38
R2	2,325	0.38
V1	1,595	0.26
V2	4	0.00
W1	114	0.02
W2	6,707	1.11
W3	974	0.16
W4	2,947	0.49
W5	39,948	6.60
XX	844	0.14
<b>Total</b>	<b>605,495</b>	<b>100</b>

These results show slightly increase for the best joining criteria – A1. For the rest the results are on the same level.

Address variables used in the different joining processes and what is returned

#### A1

Address terms in the join	Returns
Municipality number Street name House or property unit number Letter or leasehold number Post code	Coordinates and basic statistical unit

#### A2

Address terms in the join	Returns
Municipality number Modified street name House or property unit number Letter or leasehold number Post code	Coordinates and basic statistical unit

#### A3

Address terms in the join	Returns
Municipality number Modified street name	Coordinates and basic statistical unit

House or property unit number Letter or leasehold number	
---	--

#### A4

Address terms in the join	Returns
Municipality number Modified street name House or property unit number Post code	Coordinates and basic statistical unit

#### A5

Address terms in the join	Returns
Municipality number Modified street name House or property unit number Letter or leasehold number Post code	Coordinates and basic statistical unit

#### L1

This is a 2-step join – called L1<sub>a</sub> and L1<sub>b</sub> in the conceptual model above

L1 <sub>a</sub> - Variable used in the join	Returns
Organisation number	Municipality number Cadastral number Property unit number Leasehold number Post code
L1 <sub>b</sub> - Address terms in the join	Returns
Municipality number Cadastral number Property unit number Leasehold number Post code	Coordinates and basic statistical unit

#### M1

Address terms in the join - (manually coded in BR)	Returns
Municipality number Cadastral number Property unit number Leasehold number	Coordinates and basic statistical unit

#### N1

Joining description	Returns
Some industry codes in the Business Register are matched to building type in the address register and when it is only one match within one post code the coordinates and basic statistical unit are returned. Example of building types are schools, nursing homes and power stations.	Coordinates and basic statistical unit

#### R1

Joining description	Returns
For one-man companies in certain industries having the same municipality and post code as in the owners private address the coordinates and basic statistical unit are returned.	Coordinates and basic statistical unit

#### R2

Joining description	Returns
One-man companies in certain industries gets the owners private address from the Population register.	Coordinates and basic statistical unit

#### V1

Joining description	Returns
Phonetic geocoding of modified addresses in BR against Address register.	Coordinates and basic statistical unit

#### V2

Joining description	Returns
Manual coding of coordinates for Establishments around Oslo Airport while awaiting addresses for this area in the official Cadastre.	Coordinates

#### W1

Joining description	Returns
Addresses in BR joined against outdated addresses without coordinates in the Address register.	Basic statistical unit

#### W2

Joining description	Returns
Joining by municipality number and modified street name. Requires streets to be within one and only one basic statistical unit.	Basic statistical unit

#### W3

Joining description	Returns
Joining by municipality number, alternative address (name of place) and post code.	Basic statistical unit

#### W4

Joining description within BR	Returns
If the Establishment has the same municipality and post code as one or more of the Establishments joined in <b>L1</b> , the basic statistical unit as the majorities of the Establishments is returned.	Basic statistical unit

#### W5

Joining description within BR	Returns
The basic statistical unit returned is the one that have most businesses within the same post code.	Estimated basic statistical unit

**XX** - Businesses not having any kind of geocoding get XX as joining code.

### Contact information

Marianne Dysterud, Statistics Norway, [marianne.dysterud@ssb.no](mailto:marianne.dysterud@ssb.no) or [ssb@ssb.no](mailto:ssb@ssb.no)