**Mapping the church organs of Denmark**

by Søren Elkjær Trads Johansen

**Abstract**

This paper has as its stated goal to show the steps used to create an interactive map of the church organs of Denmark, on which each organ gets a point on the map containing specific information of the chosen church organ, attained from the organ database of the National Museum of Denmark. This will be done with the help of the OpenCage geocoding API to assign coordinates to the place names of the database, and then turning those coordinates into points on a map using the R package Leaflet. Alongside demonstrating a rough how-to of making the map, this paper also wishes to outline the essentials of making data open and accessible according to the FAIR-principles (primarily by showing how the National Museum of Denmark does not currently follow those principles when it comes to church organ data), as well as to highlight the importance of the concept of tidy data when working with spreadsheets as a data storage format (again by showing how the National Museum of Denmark does not work with tidy data concerning church organs).

**Keywords**

Church organs, Denmark, tidy data, FAIR principles, Microsoft Excel, spreadsheet, R, RStudio Leaflet, OpenCage, Geocoding API

**Introduction**

Denmark contains over 2200 churches as well as several institutions and private households with church organs. To systematise and make public information about this large trove of instruments, the National Museum of Denmark has undertaken the laborious project of collecting data about each of the instruments and make that data accessible online to everyone interested. This project has been going on since 2004, and is still labelled as a work in progress, with over 2700 instruments in the database as of December 2023.[[1]](#footnote-1)

This project is interesting for many reasons, chief of which is the fact that Denmark, through the public funding of its state church, has what at least one German writer on the subject calls an ideal situation for the upkeep and usage of church organs; Denmark, so to speak, is in a privileged position regarding its instruments.[[2]](#footnote-2) On a more international scale, the art of German church organ building and organ music as a whole has been declared immaterial world heritage by UNESCO in 2017[[3]](#footnote-3) (and the Germans then went on to declare the church organ the “Instrument des Jahres” in 2021[[4]](#footnote-4)), which is very relevant to Danish instruments in the sense that the Danish school of organ building has been very influenced by the (north) German school of organ building.[[5]](#footnote-5) Therefore, it is somewhat a shame when one considers the difference between the accessible online information of Danish church organs, when one compares the Danish situation to that of Germany and, especially, the Netherlands: When one takes a look at the user-driven Dutch website [www.orgelsite.nl](http://www.orgelsite.nl), one may to one’s surprise find that the Dutch seems to know and publish more about Danish church organs than the Danes themselves. This paper wishes to be a step on the way to rectify this.

**Software framework**

Before getting into the thick of the matter, this paper wants to give full disclosure about the working environment used: The work was done on a Lenovo IdeaPad S340 with 8 GB RAM, with an AMD Ryzen 7 3700U 2.30 GHz processor, running hotly, and with a 64-bit installation of Windows 10. The software used was Microsoft Excel and R version 4.3.1, utilized in the working environment of RStudio, version 2023.06.2+561.

**Data acquisition**

All the data used in the making of the map was attained from the (church) organ table on the website of the “Organ Registrant” of the National Museum of Denmark at orgel.natmus.dk. Unfortunately, it was impossible to download the entire table at once, and upon asking the National Museum of Denmark if they would be able to send the raw data in a common file format (e.g. Excel-spreadsheet or a .csv-file), they replied with a polite, but firm, no. Therefore, it was necessary to copy and paste the values of the organ table into an Excel-spreadsheet 25 values at the time, making sure that by each copy-and-paste operation no values overlapped or was missed. In practice, it took well over 2 hours to copy the entire table manually into Excel, but once done, the hard part was over. Afterwards, the raw data was cleaned thoroughly in accordance with the standards of Tidy Data in order to make it easy for both humans and machines to read.[[6]](#footnote-6) In this process, regex101.com was used to delete certain types of data pertaining to the location data of the instruments with the use of regular expressions. This was done because the location data on orgel.natmus.dk was entered with the now obsolete Danish administrative terminology, which the geocoding API would have been unable to read, e.g., the string “Ryde [Vinderup] Kirke Ringkøbing Amt” was changed to “Ryde Kirke”; the town name in brackets refer to a since 2007 defunct municipality (Danish: kommune), and the last two words refer to the also since 2007 obsolete administrative division of the county (Danish: amt).[[7]](#footnote-7)

It should be noted that while cleaning up the dataset, all the in-between-stages of data was kept organised according to the recommendations of Broman & Woo (2018).

Et billede, der indeholder tekst, kort, skærmbillede

Automatisk genereret beskrivelseSince it was only possible to push 2500 geocoding requests with OpenCage per day,[[8]](#footnote-8) it was decided to cut away all the instruments which weren’t a church instrument in the dataset. This was done within Excel by filtering out all observations not containing the words “Kirke” or “-kirke”. Doing this, the number of observations fell from 2765 to 2393 – well within the limits of the OpenCage geocoding API. The result of the arduous and laborious cleaning of the data resulted in the data found in the Excel-file “dkorgans\_tidydata\_onlykirker.xlsx” (with an accompanying .csv-file),[[9]](#footnote-9) which was used in the Rmarkdown-file to produce the map of the church organs of Denmark. Alongside the data, a readme-file was made which describes the terminology of the dataset.

Figure 1: Example from the map, with one of the three instruments of the Cathedral of Aarhus selected.

**Results**

With the map being constructed as shown in the Rmarkdown-file, the map was quickly brought up and running. Figure 1 shows that the map is working as intended, with clickable points for each instrument. In the upper right corner, there are three available map types to choose from. Really, the inclusion of Esris streetmap baselayer is superfluous since it correlates with OpenStreetMaps standard baselayer, but it was included nonetheless, because it seemed nicer to have three map types to choose from instead of two. The decision to include three map types was therefore not a scientific one, but a matter of aesthetical taste.

Et billede, der indeholder skærmbillede, tekst, kort, diagram

Automatisk genereret beskrivelse Immediately visible are some of the shortcomings of the map. If we take a look at the village church of Søllested on Lolland (figure 2), we see that this church, according to the map, has two instruments. The author of this paper, though, used to play in said church as a teenager and can confirm that Søllested on Lolland only has one organ (and a very mediocre one at that). The reason the map displays two instruments in this church is that there is another place called Søllested on the island of Funen. The map fuses these two Søllested’s into the Søllested on Lolland because they were originally listed as “Søllested [Lolland] Kirke Storstrøms Amt” and “Søllested [Fyn] Kirke Fyns Amt” on orgel.natmus.dk, respectively. For reasons explained earlier, it was necessary to trim both of those strings down to “Søllested Kirke”, or else the geocoding API would have had no chance of finding either of those locations. The geocoding API used for this map – which makes use of OpenStreetMap – thus limits the accuracy of the map slightly, since it is unable to parse the obsolete geographical terminology of the location strings, and therefore needs to be fed cleaned (and, really, dumbed-down) strings in order to find anything. Incidentally, if one searched for “Søllested [Fyn] Kirke Fyns Amt” on Google Maps, it finds the correct church right away. Put shortly, with the tools available for the construction of this map, it was either getting no map at all, or a functional, if at times a bit confused, map.

Figure 2: The map incorrectly shows the village church of Søllested, Lolland as having two church organs. This happens because the map confuses the organ in the church of the same name on Funen with the organ in the one on Lolland and conflates the two, resulting in the map thinking that Søllested on Lolland has two organs.

Et billede, der indeholder tekst, Plan, skærmbillede, diagram

Automatisk genereret beskrivelse This all goes to show that the tools at one’s disposal when working with digital humanities sometimes forces one’s hand when it comes to settling the balance between the qualitative and quantitative aspects of the data.

Another unfortunate aspect of this map – or rather, of the data as it was once entered on orgel.natmus.dk – becomes visible if we look at the church of Saint Lukas in Aarhus (figure 3). This church has a quite large organ,[[10]](#footnote-10) but it is completely invisible on the map. The reason for this is that the church was misspelled as “Sct. Lukas Kirke” instead of “Skt*.* Lukas Kirke”. This simple misspelling in the Danish abbreviation of “saint” throws the geocoding API off balance and makes it unable to locate the church, wherefore the map is unable to display the data pertaining to the otherwise nice, neobaroque instrument of this church.[[11]](#footnote-11)

Figure 3: The Church of Saint Lukas in Aarhus doesn’t have an organ according to the map, because of an unfortunate misspelling on orgel.natmus.dk.

Abovementioned drawbacks aside, it is worth noting that the map gets the majority of the instruments correctly placed on the map, and correctly provides and displays information about them. The usefulness of the map will be further discussed below.

**Critical evaluation**

The goal of this project was to assemble a map of the church organs of Denmark. This goal was accomplished, though because of its shortcomings, it should be viewed as a proof-of-concept. The most interesting part of this project has been the insight into the importance of Tidy Data on one hand, and on the other hand how the FAIR-principles (or lack thereof) made the data acquisition process an arduous one.

Et billede, der indeholder tekst, skærmbillede, Font/skrifttype, nummer/tal

Automatisk genereret beskrivelse The principles of Tidy Data by Wickham (2014) are simple, but has many implications: Put shortly, making data tidy means cleaning it so it fits into a rectangle, wherein each variable has its own column and each observation its row. In the dataset used for the map, one of the main problems was the amalgamation of several forms of information in the location column, which was necessary to clean away in order to make the geocoding API work with it. The necessity to deal away with outdated geographical terminology is just one example; another one was the sad, but necessary choice to delete otherwise interesting qualitative data about the instruments, which were mashed into the location column on orgel.natmus.dk. This was namely words specifying the type of organ (“kororgel” = choir organ, “positiv” = chair organ, “øveorgel” = practice organ, for example). Ideally, this information would have been pulled out of the location column and placed into a new column named “notes”, in order to do the Tidy Data-concept of one value per cell justice. Time, however, was of the essence, and it was decided to simply cut those otherwise interesting qualitative bits Et billede, der indeholder tekst, skærmbillede, visitkort

Automatisk genereret beskrivelseof information out of the data.

The FAIR guiding principles as specified by Wilkinson et al. (2016) came into focus in this project in the sense that the National Museum of Denmark really did not make much use of them.

Figure 4: A comparison of the technical data provided by our map, and the data on a German Wikipedia article about the very same instrument – same organ, very different visualisation of data!

(<https://de.wikipedia.org/wiki/Dom_zu_Roskilde#Orgel>, screenshot taken at 04.12.2023)

The “F”[[12]](#footnote-12) – “Findable” – can be somewhat checked off since the data was neatly and publicly visible on orgel.natmus.dk. Though, an international, standardized terminology in the world of church organs has not yet been attempted and is therefore not applied on the data on orgel.natmus.dk, which is why this project uses a home-spun terminology (as laid out in the dataset\_terminology.txt in the data folder on GitHub). The description of metadata on orgel.natmus.dk in the sections “Vejledning” and “Om Orgelregistranten” (English: “Instruction” and “About the organ registrant”, respectively) are also very short and not very thorough; the date of the observation of a given organ and other useful metadata are not included in the dataset at all, for example.

The “A” – “Accessible” – only conditionally applies to the data on orgel.natmus.dk, since it was an arduous process downloading it, and the National Museum did not want to or was not able to provide the (meta-)data per email when asked.

The “I” – “Interoperable” – does hardly apply here since the website of the dataset is entirely in the Danish language. Add to that the lack of international terminology about church organs as well as the complete lack of metadata, and you get almost no interoperability.

The “R” – “Reusable” – is slightly present. The data is, though without license, publicly available on a website belonging to a public institution, with scarce documentation about provenance.

Shortly put, regarding publicly available and easily accessible information about church organs in Denmark, there is still some way to go.

Also, it should be noted that the data on orgel.natmus.dk, and therefore our map, does not have the qualitative depth that for example a German Wikipedia article about a Danish church organ has (cf. Figure 4). However, this is of minor importance, since the goal of orgel.natmus.dk and our map is to provide an overview of the instruments, and not a qualitative in-depth description.

**Conclusion**

This project sought to create a map of all the church organs in Denmark wit the use of the data found at orgel.natmus.dk and the OpenCage geocoding API. Because – or despite of – the messiness and un-FAIR-ness of the data on orgel.natmus.dk and the deficiencies of the OpenCage geocoding API, a functional, if imperfect map was constructed. The project and its map thus not only serve as a tiny step on the way to improve the publicity and public availability of data regarding Danish church organs, but also as an example of the importance of Tidy Data and the FAIR principles, and their implications.

**References**

Broman, Karl & Woo, Kara H. (2018): ”Data Organization in Spreadsheets”; in: *The American Statistician*, 72:1, 2-10.

Danish ministry of the Interior (2023): “Kommunalreformen i 2007” (<https://im.dk/arbejdsomraader/kommunal-og-regionaloekonomi/kommunale-opgaver-og-struktur/kommunalreformen-i-2007>) (accessed 04.12.2023).

Danish Organ Index (2023): “Dansk Orgel Index” (<https://orgel.natmus.dk/>) (accessed 04.12.2023).

Faber, Rudolf & Hartmann, Philip (udg.) (2010): *Handbuch Orgelmusik – Komponisten, Werke, Interpretation*; Kassel: Bärenreiter.

German UNESCO-commission (2023): “Orgelbau und Orgelmusik als Immaterielles Kulturerbe der Menschheit anerkannt“ (<https://www.unesco.de/kultur-und-natur/immaterielles-kulturerbe/immaterielles-kulturerbe-weltweit/orgelbau-und-orgelmusik>) (accessed 04.12.2023).

OpenCage (2023): “Rate/Usage Limits” (<https://opencagedata.com/api#rate-limiting>) (accessed 04.12.2023).

The Schleswig-Holstenian State Musical Council (2021): “Instrument des Jahres” (<https://2021.instrument-des-jahres.de/>) (accessed 04.12.2023).

Wilkinson, Mark et al. (2016): “The Fair Guiding Principles for Scientific Data Management and Stewardship"; in: *Scientific Data* 3:160018.

Wickham, Hadley (2014): “Tidy data”; in: *The Journal of Statistical Software* 59:10, 1-23.

Website of the Danish labour union of conservatory-educated organists DOKS.dk (2023): “Skt. Lukas Kirke” (<https://www.doks.dk/organistbogen/aarhus-stift/sankt-lukas-kirke-aarhus#c35751>) (accessed 04.12.2023).

Wikipedia.de (2023): ”Dom zu Roskilde” (<https://de.wikipedia.org/wiki/Dom_zu_Roskilde#Orgel>) (accessed 04.12.2023).

**Required metadata**

|  |  |  |
| --- | --- | --- |
| **#** | **(executable) software metadata description** | **Software used** |
| S1 | Current software version | RStudio version 2023.06.2+561 (R version 4.3.1) |
| S2 | Permanent link to executables of this version | <https://github.com/Digital-Methods-HASS/AU672340_Johansen_Soeren/tree/main/Final%20Project> |
| S3 | Legal software license | GNU General Public License, Version 3, 29 June 2007  <https://github.com/Digital-Methods-HASS/AU672340_Johansen_Soeren/blob/main/Final%20Project/LICENSE> |
| S4 | Computing platform/operating system | Lenovo Ideapad running Windows 10 64 bit, 8 GB RAM, AMD Ryzen 7 3700U 2.30 GHz |
| S5 | If available, link to user manual - if formally  published include a  reference to the  publication in the  reference list | <https://leafletjs.com/examples/quick-start/>, accessed 05.12.2023 |
| S6 | Installation requirements & dependencies | RStudio version 2023.06.2+561 (R version 4.3.1)  Tidyverse package  Leaflet package  opencage package  dplyr package  usethis package  Microsoft Excel  GitHub |
| S7 | Support email for questions | sorenxd123@gmail.com |

**Weekly homework**

**W36: Regulære udtryk**

1. <https://regex101.com/r/B6hcAV/1>. Jeg tilføjede \b først og sidst for at undgå hypotetiske falske positiver (sæt nu at der i den fulde tekstmængde var nogle slåfejl, der gjorde at der stod nogle tal klos op ad datoangivelserne).
   1. <https://regex101.com/r/eZuPp6/1>. Her skal man dog selv lige huske at tilføje anførselstegn først og sidst i løsningen.
   2. <https://regex101.com/r/7vkTXf/1>. Her skal man dog huske at fjerne anførselstegn først og sidst i løsningen.
2. Med udgangspunkt i Broman & Woo (2018) gælder det om at gøre data i regneark læseligt for både maskiner og mennesker. Det gøres bedst ved at være konsekvent i den dataterminologi (inklusive dato-format!) man anvender i regneark (en god tommelfingerregel er at holde navnene korte, betydningsfulde og mellemrumsløse), samt at man undgår tomme celler i sit regneark. Her bør man hellere angive Ø-værdier med ”NA” eller lignende. Man kan med fordel lave et dokument ved siden af, hvor man forklarer sit regnearks dataterminologi nærmere.

På et mere overordnet plan anbefales det, at man kun har ét data pr. celle, samt at man kun har ét rektangulært datasæt pr. regneark. På den måde undgår man forvirring med at skulle navigere rundt i ét enormt regneark med data alle mulige steder – forudsat man har et konsekvent og forståeligt navnesystem til sine regneark!

Det frarådes også at lave enhver form for databehandling i sit dataregneark, eller at forsøge at visualisere data ved hjælp af typografiske eller layouttekniske hjælpemidler – disse kan nemlig ikke nødvendigvis forstås af en computer, og ville derfor kunne gå tabt i det videre databehandlingsforløb. Som helgardering kan man eventuelt gøre sit dataregneark skrivebeskyttet.

Endelig bør man tage backup af sine data og gemme det flere forskellige steder – gerne både i skyen og på et (eksternt) drev. Man har sjældent brug for de backups, man tager – lige indtil man akut har brug for dem! Det anbefales også at gemme de løbende versioner af dataregnearket i hver sin fil, i tilfælde af man får brug for at gøre nogle ændringer om (livrem-og-sikkerhedsseler-princip).

1. Regex101 syntes datasættet var for stort til at ville give mig et link – Jeg undersøgte først hvilke af de seks varianter rent faktisk var forhånden i datasættet, og fandt varianterne ”Dis Manibus”, ”Ds Manibus” og ”Diis Manibus” (alle eventuelt med ”sacrum” klasket på). Min kommando var

Di\*s\s+Manibus sacrum\b|Di\*s\s+Manibus\b

Jeg fik 20582 resultater.

**W37: OpenRefine**

<https://github.com/Digital-Methods-HASS/AU672340_Johansen_Soeren/tree/main/Final%20Project/Homework/W37%20OpenRefine>

Note the two tabs in the Excel spreadsheet – the first is the list of kings and queens of Denmark, the second are the answers to the OpenRefine questions.

**W38: Starting with R**

<https://github.com/Digital-Methods-HASS/AU672340_Johansen_Soeren/tree/main/Final%20Project/Homework/W38%20Intro%20to%20R>

The homework in of itself is stored in the RStudio document “AfleveringW38.R”.

**W39: From Danish kings to animations – make it move!**

<https://github.com/Digital-Methods-HASS/AU672340_Johansen_Soeren/tree/main/Final%20Project/Homework/W39%20Make%20it%20move>!

**W40: Maps with Leaflet**

<https://github.com/Digital-Methods-HASS/AU672340_Johansen_Soeren/tree/main/Final%20Project/Homework/W40%20Maps%20with%20Leaflet>

1. Cf. <https://orgel.natmus.dk/>, accessed 04.12.2023 [↑](#footnote-ref-1)
2. Faber & Hartmann (2010): 584. [↑](#footnote-ref-2)
3. German UNESCO-commission, <https://www.unesco.de/kultur-und-natur/immaterielles-kulturerbe/immaterielles-kulturerbe-weltweit/orgelbau-und-orgelmusik>, accessed 04.12.2023. [↑](#footnote-ref-3)
4. The Schleswig-Holsteinian State Musical Council (in collaboration with the other State Musical Councils of the Federal Republic of Germany), <https://2021.instrument-des-jahres.de/>, accessed 04.12.2023. [↑](#footnote-ref-4)
5. Faber & Hartmann (2010): 584f. [↑](#footnote-ref-5)
6. Cf. Wickham (2014): 5. [↑](#footnote-ref-6)
7. Cf. Danish ministry of the Interior, The municipal reform in 2007, <https://im.dk/arbejdsomraader/kommunal-og-regionaloekonomi/kommunale-opgaver-og-struktur/kommunalreformen-i-2007>, accessed 04.12.2023. [↑](#footnote-ref-7)
8. OpenCage rate limiting, <https://opencagedata.com/api#rate-limiting>, accessed 04.12.2023. [↑](#footnote-ref-8)
9. Attainable at <https://github.com/Digital-Methods-HASS/AU672340_Johansen_Soeren/tree/main/Final%20Project/data>, accessed 05.12.2023. [↑](#footnote-ref-9)
10. Cf. the website of the Danish labour union of conservatory-educated organists, Skt. Lukas Kirke, <https://www.doks.dk/organistbogen/aarhus-stift/sankt-lukas-kirke-aarhus#c35751>, accessed 04.12.2023. [↑](#footnote-ref-10)
11. Ibid. [↑](#footnote-ref-11)
12. The following elaborations of the FAIR acronym are all based on Wilkinson et al. (2016): 4. [↑](#footnote-ref-12)