

# GlobCovid

SDI Infrastructure for a Global Covid Dashboard



## Authors

Niklas Jaggy  
Simon Meyer

## Principal Investigators

Dr. Barbara Scherthanner-Hofer  
Dr. Manfred Mittlböck

11. Feb. 2022

# Table of Contents

- 1 Project Assignment .....1
- 2 Concept.....2
- 3 Project Management.....3
  - 3.1 Project Overview Document.....3
  - 3.2 GitLab .....4
- 4 Implementation.....5
  - 4.1 Architecture.....5
  - 4.2 SDI methods and Software .....6
- 5 Results .....10
- 6 Discussion and Outlook.....12
  - 6.1 Discussion.....12
  - 6.2 Outlook .....13
- 7 References.....15
- 8 Appendix .....16

# 1 Project Assignment

This project was assigned, planned, and carried out within the university course *Spatial Data Infrastructures: Services Implementation* at the Paris-Lodron-University Salzburg. The main project goals, besides the finished product, was to gain understanding of spatial data infrastructures, how spatial information is organized, distributed, and presented from data acquisition to presentation. These main concepts that may vary on the implementation side between organizational, enterprise, national or global levels but conceptually hold true were to be implemented at a smaller scale within this final project.

More precisely, an own SDI strategy needed to be thought of that includes the creation and sharing of standardized geoservices. With the Covid-19 pandemic and the vast amount of regularly produced data around it a general suitable application case was found, although the focus point as well as the purpose and addressed stakeholders was free to be set by the individual group. Nevertheless, it served as showcase to practice how to integrate data from various sources into a spatial data infrastructure, to prepare geoinformation in a way it becomes useful (data modeling, geocoding, quality assessment, ...), develop standardized geoservices, create descriptive metadata and to communicate information using GI services. The key tasks for this project are shown in the following:

1. Identification and development of reproducible workflows for combining and integrating identified geospatial and additional data sources.
2. Organization of geospatial information in ISO compliant manner in geodatabases with ISO compliant metadata descriptions.
3. Create and publish standardized services to share out the resources.
4. Present and communicate the services in an adequate format.

After the project topic was identified, the project management process began in which the project work was divided and each step elaborated, summarized within a project overview document. This document was created for presentation to the principal investigators (B. Schernthanner-Hofer, M. Mittlböck) from which the approval was necessary to simulate getting the contract and to kick off the project.

## 2 Concept

### *First Draft*

The original idea for the project concept was to visualize a global status report on the Corona pandemic in addition with the status of international vaccination campaigns such as COVAX. The result would have been able to show global disparities in the degree of how affected a country is, the current status of national vaccination campaigns combined with the global flow of vaccination aid in the form of international aid. However, a first search for suitable data that would be accessible and could be managed to be integrated in the SDI was not found. Although the typical indicators of the pandemic are collected and distributed by different sources, mostly depending on the administrative level, data on the vaccine rollout by initiatives such as COVAX were sparse, if at all accessible. The integration was therefore declared not feasible, and other ideas were collected.

### *Final Project Concept*

The topic on which the project in the end was narrowed down shifted away from the global vaccine campaign efforts towards the integration of information on global socioeconomic attributes at country level. The idea was to illustrate the daily global Corona status in the light of the developing status of a country as this usually relates to the level of healthcare. This adds new dimensions to the understanding of the global status and for the comparison of disparities between countries, regions or continents.

While the source of the Corona data was already secured, socioeconomic data was taken from the data center of the United Nations development program (UNDP). Combined with a geospatial dataset of global administrative boundaries all relevant data was available to realize the project concept.

### 3 Project Management

The entire project was planned as an effort spanning multiple months and constituted many individual steps, divided into different project phases. This meant that a thorough project management was necessary to assure a successful outcome.

#### 3.1 Project Overview Document

Group work and the long duration of the project made it necessary to start with the documentation of the project early on. Already during the initial conceptualization phase of the project, important decisions were documented. The documentation and management of the project was done by creating a structured project overview document in which key work packages were outlined and milestones were defined. Here, every important information regarding the project was recorded. The project overview document contained the most important information about the project itself, the projects content and goals, the frame of the project regarding resources and budget, the break-down and overview of the project structure with insights into the individual work packages, the work plan, the milestone plan, and overview graphics including a GANTT chart and risk matrix.

The project management document was accordingly updated throughout the development of the project with the adjustment of timeframes and milestones as well as the completion rate of individual work packages and the risk assessment.

The project work itself was structured into five main sections: project management, project conceptualization, setup of necessary backend services and data storage (automated data requesting and database setup), dashboard implementation, and finally project documentation, presentation and evaluation. This structure was the foundation of the work packages that were created to assist in the project's development (see Figure 1).

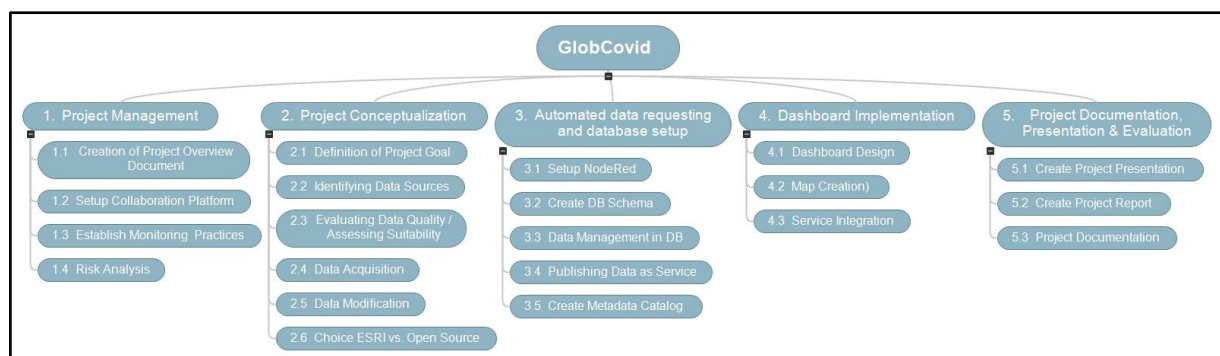


Figure 1: Work packages of the project.

The individual work packages were again subdivided into separate tasks. Some of them were active during the whole project, such as the documentation, while others were of shorter duration. The duration of these tasks is best described in the Gantt chart. In this chart current developments were documented and visualized, most importantly the time and duration of tasks (with start and end dates) as well as their completion rate. The final chart can be seen in Figure 2.

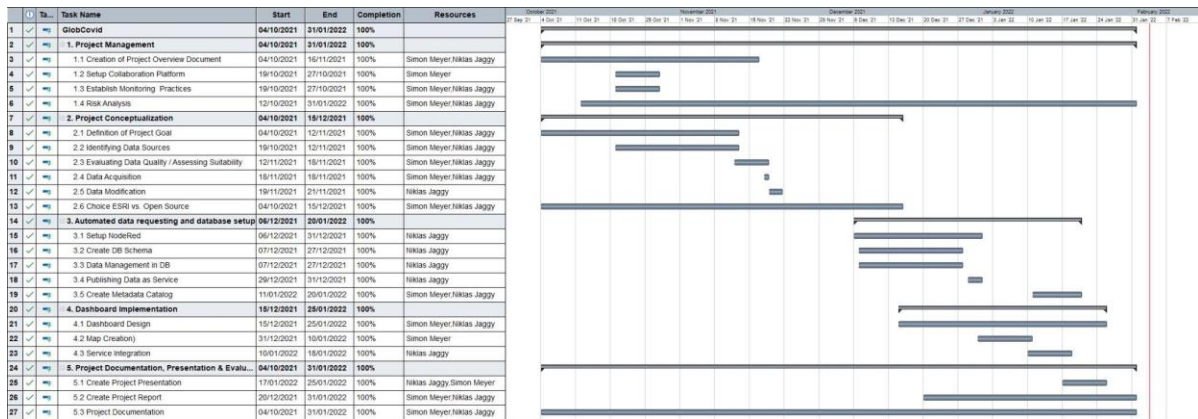


Figure 2: Overview of the Gantt chart created for the project management.

The final document, signed and approved by the principal investigators, served as guideline for splitting work as well as for time management throughout the entire duration of the project and ensured a structured work process due to the provided framework.

### 3.2 GitLab

Complementary to the project overview document a new GitLab project was set up. GitLab is a web-platform by GitLab Inc. that allows users to develop, manage and operate their software in groups. GitLab itself is developed as an open-source project which makes it an appealing option especially for smaller developers (GitLab, 2022).

In this project GitLab was mainly used as a central platform to share new data sources between the developers, as a central deposit of important datasets and documents, and as a site where developmental steps could be documented. For this, three new directories were created. The first was designated to contain main data files, such as the country boundaries dataset, as well as provisional files for testing. The second directory was created to contain the documentation and the third directory was created to include metadata information of the files used in the project. Besides the main directories, a readme document was created which gives an overview of the project's working title, the team members and the most important links to other platforms linked to the project. The most important links were the link to the current version of the dashboard, the link to the project overview document (hosted as Google Docs file), the link to the main Our World In Data Covid-19 data repository on GitHub, the link to the HDI dataset source, the link to the source of the administrative boundaries dataset (NaturalEarthData), and links to important services related to the project's development such as the Feature Server and the Web Feature Server. For a structured overview on the documentation (and implementation) of the dashboard a wiki page was created as well, where important steps could be documented.

## 4 Implementation

With the project management framework at hand, the implementation of the project, meaning the subsequent realization of each work package began. This meant to implement the SDI methods and tools that we were acquainted with throughout the course in an SDI architecture of our own development. The following sections will first outline the SDI architecture that constitutes this project while the following section describes the various SDI methods and tools applied in the architecture.

### 4.1 Architecture

This section is dedicated to outline and describe the architecture of the spatial data infrastructure therefore serving as framework for the project implementation. An overview can be found in Figure 3. According to how spatial data infrastructures are defined, they construct a framework containing geoinformation, metadata, tools and the users. This framework then allows to acquire, process, use, distribute and maintain spatial data, which are all processes that at their foundation require a set of technologies, standards, and policies to realize them. In the following, the structure on how the technologies and standards are laid out and how they relate to each other within the project is documented.

The starting point for every spatial data infrastructure is the actual data, where it comes from, what needs to be done with it and what its purpose is. For this project, data from three different sources were taken to be combined in a single dataset. This dataset, containing the geoinformation, socioeconomic attributes and the recent Covid data, is then cleaned, shared out and within a web map embedded in a Dashboard. While the geospatial and socioeconomic data are of a static kind, the Corona data is requested on a daily basis. This event-driven request is performed by NodeRed, which after grabbing data also processes it and injects it into a PostgreSQL geodatabase.

The geodatabase is used as data storage for all project data. The socioeconomic enriched geospatial data as well as the daily grabbed Covid data is collected in the database in the form of tables and then combined in a single dataset. At this point, the data is safely stored but cannot be used in any kind of way. Therefore, it is necessary to share the data over the web as a service to make it integratable. For this project an ArcGIS Server is used to create an OGC Web Feature Service which makes it possible to include the data in web maps. Map creation took place either in ArcGIS Online or the desktop solution ArcGIS Pro but in both cases the result needed to be a web map as it is included as a central part in the dashboard and because all relevant data is contained within the web map layers. As it is in professional scenarios, it is required to provide a comprehensive description of the data in the form of metadata. This is crucial in terms of transparency and credibility as it allows users to comprehend the content, data origins and, if occurred, data processing/altering steps. Within the project SDI, a metadata catalog was created externally following ISO standard 19139 which was then attached to the dataset.

The dashboard constitutes the endpoint of the SDI and the final output product. It is based on the web maps and the contained data which is displayed in the form of lists, indicators and graphs to fulfill the dashboard’s purpose of communicating the information as precisely and concisely as possible.

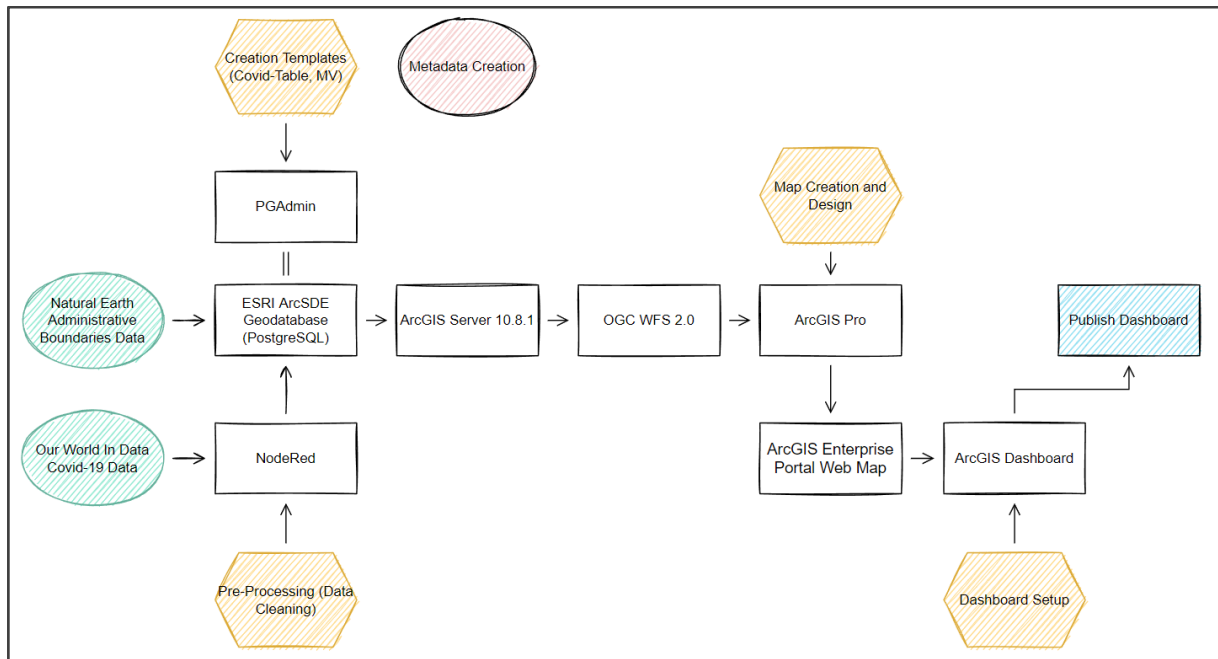


Figure 3: Overview of the project’s architecture and services.

## 4.2 SDI methods and Software

The various methods and software used to realize this project can be divided into two categories. One is the class of custom-of-the-shelf (COTS) solutions which refers to closed developed software products that need a paid license to use. On the opposite side are free and/or open-source solutions (FOSS) which are openly developed and can be used free of charge. Software and methods from both worlds are used throughout the process of this project. The primary reason for relying on COTS solutions to this degree was uncertainty regarding the data availability and the time costs of finding appropriate data. Because using solely FOSS methods, especially the development of the dashboard would have cost enormous time resources. Together with the uncertainty about the data, it was decided that including COTS solutions would be the safer path for a successful project.

Table 1: COTS and FOSS used in the project.

	Method	Functionality
<b>COTS</b>	ESRI ArcSDE Geodatabase (PostgreSQL)	Data Storage



	ArcGIS Server	Provide Web Services
	ArcGIS Map Viewer/ArcGIS Pro	Web Map Production
	ArcGIS Dashboard	Visualization Product
<b>FOSS</b>	NodeRed	Grab Daily Data
	pgAdmin	Graphical DB Management
	GeoNetwork	Metadata Creation

The following sections will provide a more detailed overview on what the individual methods are and how they are used within the project's architecture. The first one includes all methods falling into the COTS category while the latter one covers FOSS solutions.

### *COTS solutions*

Within the COTS solutions all products used are from the ESRI ecosystem. This allows an easy and convenient integration of services within a system of a single vendor. The respective software solutions are all licensed through the University of Salzburg where most of the software, apart from local ArcGIS installations, is also hosted.

### **ESRI ArcSDE Geodatabase**

The relational database used for this project needed to be able to not only handle generic data in tables but to also hold geoinformation. For this case a geodatabase was necessary. Within the ESRI ecosystem, the ArcSDE (=Spatial Database Engine) provides the capabilities to extend a relational database to enable spatial data handling. It supports many commercial database models such as Oracle and SQL Server but nowadays also the open-source database PostgreSQL (ESRI, 2022a), which is used for this database.

For the project, the PostgreSQL database served as a collection point of all datasets included. The global boundaries, enriched with socioeconomic attributes, were loaded into the database manually through ArcGIS Pro. Using the free graphical management tool for PostgreSQL databases called pgAdmin, a template table was created in which the daily Covid data is regularly injected. Next to the regular tables, a materialized view is created that joins the geospatial database and the Covid data. The advantage of the materialized view against a regular view is the fact that it is created once and not each time it is requested, therefore significantly improving performance. This materialized view, registered with the geodatabase, constitutes the output to be used in further steps of the architecture.

### **ArcGIS Server 10.8.1**

The ArcGIS server comprises the back-end server software component of ArcGIS Enterprise. It allows users to easily share out geoinformation to their own organization or in general via Web-Services such as WFS, WMS, Feature Access, etc. (ESRI, 2022b). For this project, the university hosted server was used to share out the materialized view as OGC WFS 2.0. Setting up the service is done through a server manager user interface which generates a URL to access the data. With the web service available all relevant data could be easily integrated in web maps.

### **ArcGIS Map Viewer/ArcGIS Pro**

The ArcGIS Map Viewer and ArcGIS Pro are two different options of producing and publishing web maps. While the map viewer is integrated and available in ArcGIS Enterprise (in our case in the Z\_GIS ArcGIS Enterprise Portal (Version 10.9)), ArcGIS Pro constitutes the desktop GIS solution for map production. Within the map viewer, maps are quickly drawn up and directly available as web maps, although map making functionalities are limited. On the other hand, ArcGIS Pro provides a more sophisticated environment but needs the additional step of sharing the map specifically as a web map. In this project, both methods were applied to generate a total of four maps by integrating the data through the WFS that was setup before. These maps portray the global new cases, vaccinations, boosters, and tests with an additional No Data layer below in case data for some countries are not available at the given day.

### **ArcGIS Dashboard**

The ArcGIS Dashboard app is integrated as a ready-to-use app within the university's Geoportal. Its primary goal is to serve as the visualization and information communication interface for the target audience (ESRI, 2022c). A flexible drag-and-drop mechanism allows to integrate various maps, indicators, charts, and other interactive elements to lay out the dashboard framework. Concerning the data needed for charts and indicators, all relevant information was included in the Covid dataset and therefore also updated daily. A more detailed description of the project dashboard is found in the results section.

### *FOSS solutions*

#### **NodeRed 2.x**

NodeRed is a programming tool that provides a flow-based programming setting for event-driven applications using a modular-kit system. It allows graphical development of programming flows in a browser-based editor with simple deployment of said flows. Next to the existing base package, many different nodes, such as a node allowing to connect to PostgreSQL databases, provided by an active community can be integrated in the flow. Additionally, functionality can be added via writing own JavaScript functions (Node-RED, 2022).

Within this project, NodeRed played a critical role as it allowed to easily grab the daily Covid data from the web, clean it, inject it into the geodatabase and produce the final materialized view that is used in the dashboard (see Figure 4). The event-driven part is that the data is grabbed each day as a csv file at a specified time through a http request. This csv data is then

parsed row by row and cleaned using a JavaScript code snippet. Particularly handling empty cells and some string handling is done in this step. Finally, the SQL statements are created that are needed for the data injection. Sequentially, the template is cleaned, data is injected and lastly the materialized view is refreshed using the constructed SQL statements. Once this flow was created and deployed, this process takes place on a regular basis.

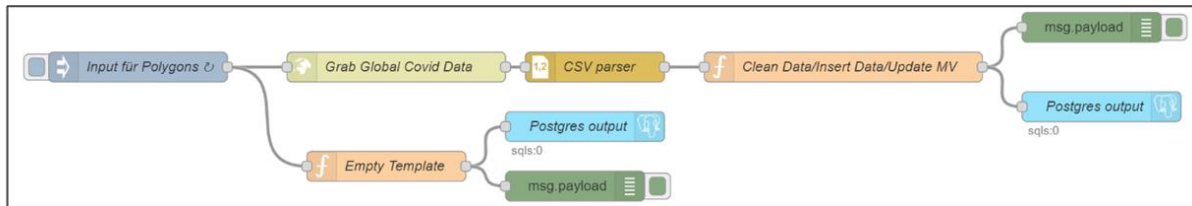


Figure 4: NodeRed flow used to grab the daily Covid-19 data.

## pgAdmin

The pgAdmin software is a tool for developing and administering PostgreSQL databases (pgAdmin, 2022). Like the open-source and free database pgAdmin is also free to use. It provides a graphical user interface for easy database management and can be extended in functionality for example for handling spatial data with the PostGIS extension. Within this project the software tool was used to simplify the management of the PostgreSQL geodatabase. Through the query tool on the user interface the Covid data template table was created as well as the materialized view. Additionally, the overall management of the geospatial dataset was also done with pgAdmin.

## GeoNetwork

GeoNetwork as software tool is a metadata information system software for managing geodata, more precisely it is a “catalog application to manage spatially referenced resources” (GeoNetwork, 2022). Of particular interest within the context of this project was the metadata editor to create and manage metadata information. This editor supports common ISO metadata standards (e.g., ISO19115) as well as the Dublin Core format. This project used it to generate one metadata catalogue for the final dataset. Therefore, it includes all information on the dataset combined within the final one as well as description on how the final dataset came to be. This metadata catalogue was then attached to the dataset using ArcGIS Pro.

## 5 Results

The final result of the SDI project and preceding workflows was the dashboard (see Figure 5). As the goal of the project was to incorporate and compare Covid-19 data and socioeconomic data, the focus of the dashboard design was set goals. The dashboard incorporates the data sources previously described and presents them in a clear structure. The intention was to make the information easily accessible as the set target audience does not consist exclusively of geospatial experts.

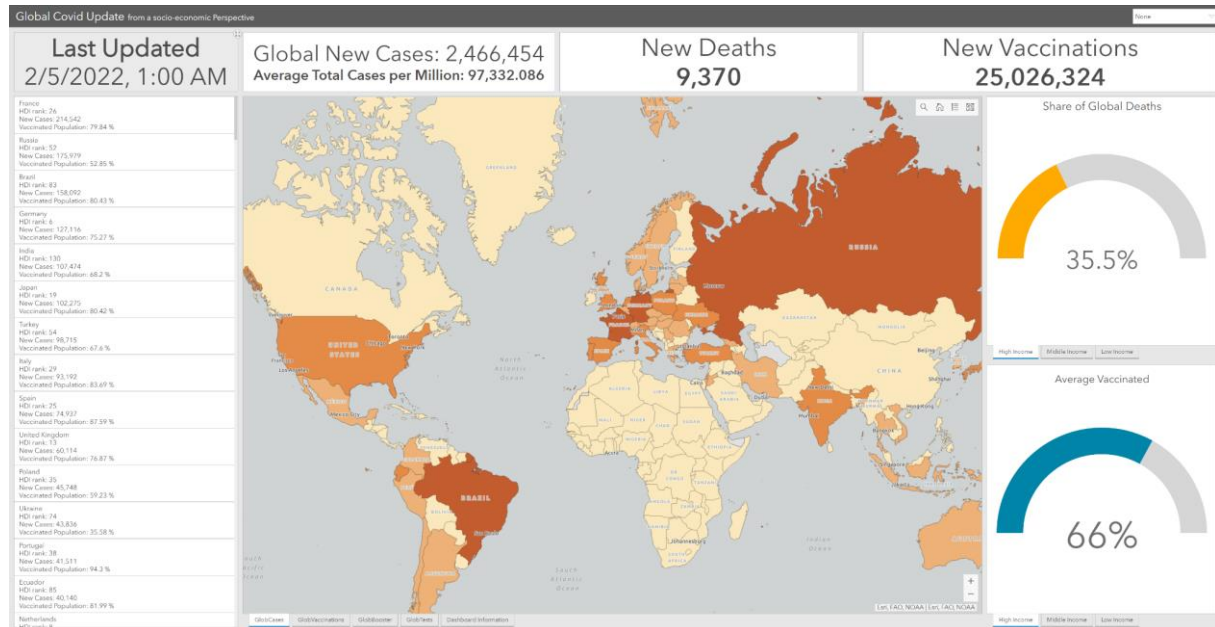


Figure 5: The dashboard in its final appearance. The central element, positioned in the center, is the map.

The result of the dashboard design and the geoservices integration was a user interface that is structured into four main sections:

1. The map frame, as the main information media, takes up the biggest section in the center of the dashboard. Altogether four maps have been created, each of them has a different focus: The first map, and the default map that is loaded when accessing the dashboard, simply shows the spatial distribution of new Covid-19 cases per country. The second map shows the number of new cases as differently sized circles in comparison to the percentage of fully vaccinated citizens per country. In addition to these standard Covid-19 statistics, two more maps were included which focus on the usage and distribution of booster vaccinations and tests per country (divided by population) in comparison to the respective HDI rank. Every map has basic map controls integrated allowing users to navigate and select/deselect individual layers of the web maps. The map's legend can be expanded by clicking on the respective control button on the top-right of the map.
2. On the left side, a country list is displayed that shows countries ordered by the number of new cases and additional information about HDI rank and the percentage of vaccinated citizens.

3. On the right side, statistics of income classes are displayed regarding the share of global deaths and percentage of vaccinations. Income classes are divided into high, medium, and low income.
4. On the top, four small boxes are displayed showing the date and time of the last update of the dashboard, the count of global new cases with the average total cases per million, the numbers of new cases and new vaccinations worldwide.

To make better use of the dynamic capabilities of dashboards, different forms of interactivity were integrated. For example, users can click on individual countries in the country list on the left side of the dashboard, then the respective country's borders will be highlighted and a zoom onto the selected country will take place. If the user wants to get further information about certain countries, a mouse click on a country will open a pop-up window showing more detailed information about current Covid-19 numbers and the socioeconomic status of the selected country. Should the user want to filter the countries based on their average income, they can do so by selecting the income group in the list on the upper right of the dashboard. Per default no filtering is applied, but selecting a filter (e.g. "High income: OECD") will have an impact on the countries in the county list on the left side of the dashboard, as well as the number of new deaths and new vaccinations displayed at the top of the dashboard. The values are then updated according to the selection made.

Information about the dashboard, its authors and the used data references were integrated into a separate window that can be selected next to the map selection. It then replaces the web map in the map frame.

# 6 Discussion and Outlook

## 6.1 Discussion

The spatial data infrastructure in its final state successfully implements all components of a typical SDI such as common technologies and standards by acquiring (spatial) data in an event-driven manner, storing it, publishing it through web services and integrating it in a concise visualization product for maximum information communication. The critical issue of bringing together the individual components, linking them and making them work together as an infrastructure was solved, resulting in a presentable final product.

From the beginning until the project finalization, the first task of setting up a proper project management provided a framework, which is certainly to be evaluated positive as it gave structure and orientation for the project development. Whilst the level of detail in terms of setting precise dates is definitely important for general project management, this project benefitted the most by the setting up an overarching structure with work packages, milestones and an estimated time frame.

As defined within the project management's risk matrix, a critical issue was finding a suitable data source on which the project could be built. This turned out to be more difficult than expected and proved to be the biggest entry hurdle, even leading to adjusting the project topic. Major issues were to find a data source that fulfills all requirements of global coverage, regularly updated (e.g., daily) and being accessible via an API or requestable through NodeRed.

As it is common with data-driven tasks, data preparation and data cleaning is essential for making data useful. The biggest issue for further processing the data grabbed from GitHub was handling incomplete information in the form of empty cells. These needed to be filled for the rows to be injected properly into the PostgreSQL table template. Other necessary corrections which included dealing with exceptional ISO codes and correcting for special characters finally allowed data usage and highlights the relevance of efforts spent in the data processing step of a project.

Related to data issues is the aspect of data volume and the consequences for the performance of the final dashboard. If the tables get too big, web-services would be slowed down which results in slower interaction at the dashboard or complete failure of presenting the data. One mitigation strategy is the conversion of the spatial data from polygons to points, thereby reducing data size. For this project, the overall table size was of limited size which allowed to stay with polygon representations without any drawbacks concerning performance as long as all systems and servers were up and running.

In conclusion, the project and its outcomes, the dashboard and working SDI, measured against the goals defined in the project management can be declared a success. Risks and critical waypoints were solved across the work packages. Remaining risks are on one side the continuous supply of data and ensuring that the underlying technology, the systems and

servers hosted by the university, keeps running. Both of these risks, however, are out of the hands of the project team.

## 6.2 Outlook

The current dashboard version communicates information on the daily global Covid situation with respect to the socioeconomic development status on country level. Within the project team several possible ways of further developing the project regarding different aspects were proposed and discussed. Paths of development include (multitemporal) data, presentation functionality and technology.

Concerning the data aspect, so far only a specific selection of socioeconomic indicators is included in the final dataset including the HDI index itself which summarizes a number of development dimensions through several indicators (life expectancy, years of schooling, GNI per capita, etc.). As there are plenty of other indicators available, a next step could be to change or extend the dataset to include a different indicator set, potentially focused on a specific dimension such as healthcare or education, or just a larger number of indicators. However, this contains the danger of ending up with a too bloated dataset which becomes incomprehensible and too overwhelming for a concise information communication.

A second consideration regarding the data topic for further development is including a temporal component in the data, meaning a days-to-day/day-to-week comparison. Currently, the dashboard shows a daily status report which prohibits displaying any developments or trends along a temporal axis. While the static, but temporal contextless data communicates the most urgent information, the comparison to past situations in form of a multitemporal dataset would allow the user to put the information in the right context, therefore gaining additional information. However, within this project it is so far not yet solved how such a multitemporal table/dataset would be created as data access is only possible for the latest day.

Following the data aspect, further development regarding the final presentation result is possible. In its current state, the dashboard allows the user to interactively explore the data per country or per income group. This also means that the interactivity potential of the dashboard app is not yet fully exploited. A next step would be to implement category filters with broader effects by updating several dashboard panels simultaneously. This would allow the visualization of the connections between the Covid situation and socioeconomic status on a broader level.

Finally, continuing development could also affect the underlying technological solutions as the present SDI is largely dependent on commercial software products. Although it was appropriate within this project and is suitable for many professional projects, free and open-source options are available for each SDI component. For example, in the form of Geoserver to host data and publish web services, QGIS to produce and publish web maps and Apache Superset as a free dashboard alternative. It must be noted that this comes at a price which usually is more time investment and higher programming knowledge to set up the systems. On the other side, the benefit would be a more transparent, reproducible, and sustainable

(regarding licensing) spatial data infrastructure. This leaves an up to this point, within the course's requirement, successfully finished project open for further development and improvement.



## 7 References

### Link to the GitLab Project:

<https://git.sbg.ac.at/s1039004/sdi-covid-19-dashboard>

### Data

NATURALEARTH (2022) ADMIN 0 – COUNTRIES. AVAILABLE AT: [HTTPS://WWW.NATURALEARTHDATA.COM/DOWNLOADS/50M-CULTURAL-VECTORS/50M-ADMIN-0-COUNTRIES-2/](https://www.naturalearthdata.com/downloads/50m-cultural-vectors/50m-admin-0-countries-2/) (ACCESSED: 11.02.2022).

OUR WORLD IN DATA (2022) COVID-19 DATA LATEST. AVAILABLE AT: [HTTPS://GITHUB.COM/OWID/COVID-19-DATA/BLOB/MASTER/PUBLIC/DATA/LATEST/OWID-COVID-LATEST.CSV](https://github.com/owid/covid-19-data/blob/master/public/data/latest/owid-covid-latest.csv) (ACCESSED: 11.02.2022).

UNITED NATIONS (2022) HUMAN DEVELOPMENT INDEX AVAILABLE AT: [HTTP://HDR.UNDP.ORG/SITES/DEFAULT/FILES/2020\\_STATISTICAL\\_ANNEX\\_TABLE\\_1.XLSX](http://hdr.undp.org/sites/default/files/2020_statistical_annex_table_1.xlsx) (ACCESSED: 11.02.2022).

### Literature

ESRI (2022A) TYPES OF GEODATABASES. AVAILABLE AT: [HTTPS://DESKTOP.ARCGIS.COM/DE/ARCMAP/LATEST/MANAGE-DATA/GEODATABASES/TYPES-OF-GEODATABASES.HTM](https://desktop.arcgis.com/de/arcmap/latest/manage-data/geodatabases/types-of-geodatabases.htm) (ACCESSED: 05.02.2022).

ESRI (2022B) WHAT IS ARCGIS SERVER? AVAILABLE AT: [HTTPS://ENTERPRISE.ARCGIS.COM/DE/SERVER/LATEST/GET-STARTED/WINDOWS/WHAT-IS-ARCGIS-FOR-SERVER-.HTM](https://enterprise.arcgis.com/de/server/latest/get-started/windows/what-is-arcgis-for-server-.htm) (ACCESSED: 06.02.2022).

ESRI (2022C) ARCGIS DASHBOARDS. AVAILABLE AT: [HTTPS://WWW.ESRI.COM/EN-US/ARCGIS/PRODUCTS/ARCGIS-DASHBOARDS/OVERVIEW](https://www.esri.com/en-us/arcgis/products/arcgis-dashboards/overview) (ACCESSED: 04.02.2022).

GEONETWORK (2022) GEONETWORK. OPEN-SOURCE. AVAILABLE AT: [HTTPS://GEONETWORK-OPENSOURCE.ORG/](https://geonetwork-opensource.org/) (ACCESSED: 06.02.2022).

GITLAB (2022) WHAT IS GITLAB? AVAILABLE AT: [HTTPS://ABOUT.GITLAB.COM/WHAT-IS-GITLAB/](https://about.gitlab.com/what-is-gitlab/) (ACCESSED: 07.02.2022).

NODE-RED (2022) NODE-RED. LOW-CODE PROGRAMMING FOR EVENT-DRIVEN APPLICATIONS. AVAILABLE AT: [HTTPS://NODERED.ORG/](https://nodered.org/) (ACCESSED: 05.02.2022).

PGADMIN (2022) PGADMIN. POSTGRESQL TOOLS. AVAILABLE AT: [HTTPS://WWW.PGADMIN.ORG/](https://www.pgadmin.org/) (ACCESSED: 04.02.2022).

UNITED NATIONS (2022) HUMAN DEVELOPMENT INDEX (HDI). AVAILABLE AT: <https://hdr.undp.org/en/content/human-development-index-hdi> (ACCESSED: 03.02.2022).

## 8 Appendix

### SQL Code

```
-- Creating table template for Covid data injection
CREATE TABLE test_template_globCovid (
  id SERIAL,
  iso_code character varying(5),
  location character varying (20),
  continent character varying(15),
  last_updated_date date,
  total_cases bigint,
  total_cases_per_million real,
  new_cases integer,
  total_deaths bigint,
  total_deaths_per_million real,
  new_deaths integer,
  total_tests bigint,
  people_fully_vaccinated bigint,
  total_boosters integer,
  new_vaccinations integer,
  people_vaccinated_per_hundred real,
  population_density real,
  median_age real,
  extreme_poverty real,
  life_expectancy real,
  human_development_index real
);
ALTER TABLE s1079105.s1079105_test_template_globCovid OWNER to
owner;
```

```
-----
--Empty table before data injection
DELETE FROM test_template_globCovid;
-----
--Create table index based on ISO codes
CREATE INDEX test_template_globCovid_idx ON test_template_globCovid
(iso_code);

CREATE UNIQUE INDEX ne_admin50m_hdi_pgtype_idx ON
ne_admin50m_hdi_pgtype(adm0_a3);
```

```

-- Create materialized view and refreshing it
CREATE MATERIALIZED VIEW globcovid_admin_points_mv
AS
SELECT d.objectid, c.iso_code, c.location,d.sovereignty,d.type,
d.subregion, c.continent, c.last_updated_date, c.total_cases,
c.total_cases_per_million, c.new_cases, c.total_deaths,
c.total_deaths_per_million, c.new_deaths, c.total_tests,
c.people_fully_vaccinated, c.total_boosters, c.new_vaccinations,
c.people_vaccinated_per_hundred, d.economy, d.gross_nati,
d.income_grp, d.pop_est, c.population_density,
c.human_development_index, d.hdi_rank, d.expected_y,
d.mean_years,c.median_age, c.extreme_poverty, d.shape
FROM ne_admin50m_hdi_points_pgtype d
LEFT JOIN test_template_globCovid c ON d.adm0_a3 = c.iso_code
WHERE iso_code IS NOT NULL;

-----
-- create table index and spatial index
CREATE UNIQUE INDEX globcovid_admin_mv_idx ON globcovid_admin_mv
(objectid);

CREATE INDEX globcovid_admin_mv_geom_idx
ON globcovid_admin_mv
USING GIST (shape);

```

## JavaScript Code

```

if ((msg.payload.iso_code.length <= 3 ||
msg.payload.iso_code.includes("KOS")) &&
msg.payload.iso_code.length>0) {
  var isoCode = "";
  if (msg.payload.iso_code != null){
    if (msg.payload.iso_code.includes("KOS")){
      isoCode="KOS";
    }else{
      isoCode = msg.payload.iso_code;
    }
  }

  }else{}

  var location = "";
  if (msg.payload.location != null){

```

```
        if (msg.payload.location.includes("")){
            location =
msg.payload.location.replace(msg.payload.location, "Ivory Coast");

        }else{
            location = msg.payload.location;
        }
    }

    var continent = "";
    if (msg.payload.continent != null){
        continent = msg.payload.continent;
    }else{}

    var lastUpdated = "0000-00-00";
    if (msg.payload.last_updated_date != null){
        lastUpdated = msg.payload.last_updated_date;
    }else{}

    var totalCases = 0;
    if (msg.payload.total_cases != null){
        totalCases = msg.payload.total_cases;
    }else{}

    var totalCasesMillion = 0;
    if (msg.payload.total_cases_per_million != null){
        totalCasesMillion = msg.payload.total_cases_per_million;
    }else{}

    var newCases = 0;
    if (msg.payload.new_cases != null){
        newCases = msg.payload.new_cases;
    }else{}

    var totalDeaths = 0
    if (msg.payload.total_deaths != null){
        totalDeaths = msg.payload.total_deaths;
    }else{}

    var totalDeathsMillion = 0;
    if (msg.payload.total_deaths_per_million != null){
        totalDeathsMillion = msg.payload.total_deaths_per_million;
    }else{}
}
```

```
var newDeaths = 0;
if (msg.payload.new_deaths != null){
    newDeaths = msg.payload.new_deaths;
}else{}

var totalTests = 0;
if (msg.payload.total_tests != null){
    totalTests = msg.payload.total_tests;
}else{}

var peopleVaccinated = 0;
if (msg.payload.people_fully_vaccinated != null){
    peopleVaccinated = msg.payload.people_fully_vaccinated;
}else{}

var totalBoosters = 0;
if (msg.payload.total_boosters != null){
    totalBoosters = msg.payload.total_boosters;
}else{}

var newVaccinations = 0;
if (msg.payload.new_vaccinations != null){
    newVaccinations = msg.payload.new_vaccinations;
}else{}

var vaccinatedHundred = 0;
if (msg.payload.people_vaccinated_per_hundred != null){
    vaccinatedHundred =
msg.payload.people_vaccinated_per_hundred;
}else{}

var popDensity = 0;
if (msg.payload.population_density != null){
    popDensity = msg.payload.population_density;
}else{}

var medianAge = 0;
if (msg.payload.median_age != null){
    medianAge = msg.payload.median_age;
}else{}

var extremePoverty = 0;
```

```

if (msg.payload.extreme_poverty != null){
    extremePoverty = msg.payload.extreme_poverty;
}else{}

var liveExpectancy = 0;
if (msg.payload.life_expectancy != null){
    liveExpectancy = msg.payload.life_expectancy;
}else{}

var hdi = 0;
if (msg.payload.human_development_index != null){
    hdi = msg.payload.human_development_index;
}else{}

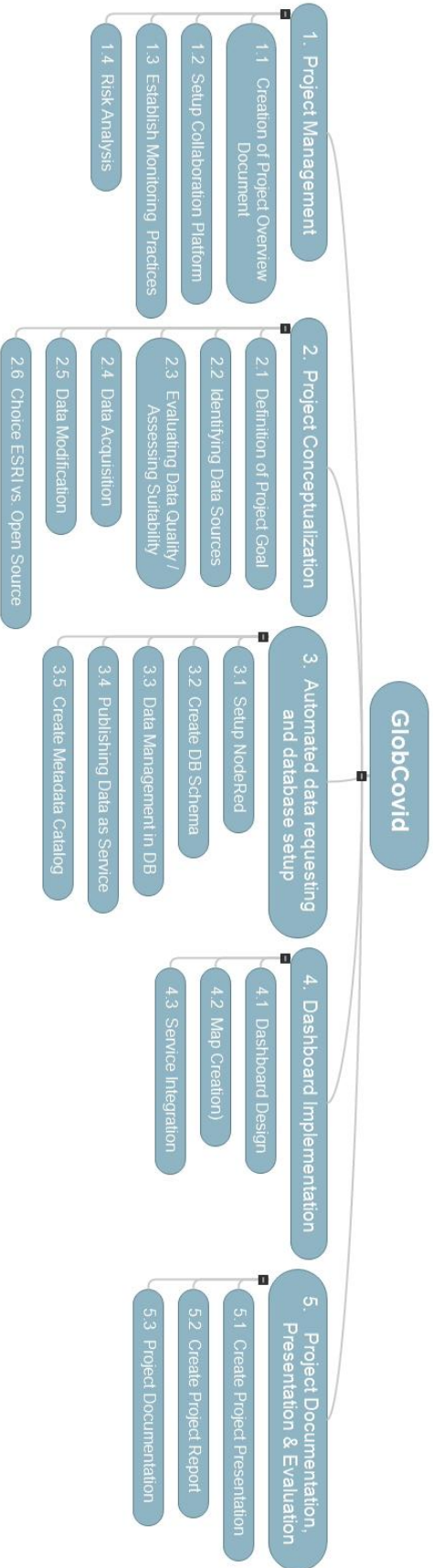
var insertStatement = "";
insertStatement += "INSERT INTO test_template_globCovid
(iso_code,location,continent,last_updated_date,total_cases,total_ca
ses_per_million,new_cases,total_deaths,total_deaths_per_million,new
_deaths,total_tests,people_fully_vaccinated,total_boosters,new_vacc
inations,people_vaccinated_per_hundred,population_density,median_ag
e,extreme_poverty,life_expectancy,human_development_index) VALUES"
+ "("+ "'" + isoCode + "','" + location+ "','" + continent + "','"
+lastUpdated+ "','" +totalCases+ "','" +totalCasesMillion+ "','"
+newCases+ "','" +totalDeaths+ "','" +totalDeathsMillion+ "','"
+newDeaths+ "','" +
totalTests+ "','" +peopleVaccinated+ "','" +totalBoosters+ "','" +newVacc
inations+ "','" +vaccinatedHundred+ "','" +popDensity+ "','" +medianAge+
',''+extremePoverty+ "','" +liveExpectancy+ "','" +hdi+''");";

msg.payload = insertStatement;
return msg;
} else {
}

var updateMviewStatement = "";
updateMviewStatement = "REFRESH MATERIALIZED VIEW CONCURRENTLY
globcovid_admin_mv;";
msg.payload = updateMviewStatement;
return msg;

```

# Gantt Chart & Project Work Packages



ID	Task Name	Start	End	Completion	Resources
1	<b>1. Project Management</b>	04/10/2021	31/01/2022	100%	
2	1.1 Creation of Project Overview Document	04/10/2021	31/01/2021	100%	Simon Meyer,Niklas Jaggy
3	1.2 Setup Collaboration Platform	19/10/2021	27/10/2021	100%	Simon Meyer
4	1.3 Establish Monitoring Practices	19/10/2021	27/10/2021	100%	Simon Meyer,Niklas Jaggy
5	1.4 Risk Analysis	12/10/2021	31/01/2022	100%	Simon Meyer,Niklas Jaggy
6	<b>2. Project Conceptualization</b>	04/10/2021	19/12/2021	100%	
7	2.1 Definition of Project Goal	04/10/2021	12/11/2021	100%	Simon Meyer,Niklas Jaggy
8	2.2 Identifying Data sources	19/10/2021	12/11/2021	100%	Simon Meyer,Niklas Jaggy
9	2.3 Evaluating Data Quality / Assessing Suitability	12/11/2021	18/11/2021	100%	Simon Meyer,Niklas Jaggy
10	2.4 Data Acquisition	18/11/2021	18/11/2021	100%	Simon Meyer,Niklas Jaggy
11	2.5 Data Modification	19/11/2021	21/11/2021	100%	Niklas Jaggy
12	2.6 Choice ESRI vs. Open Source	04/10/2021	19/12/2021	100%	Simon Meyer,Niklas Jaggy
13	<b>3. Automated data requesting and database setup</b>	08/12/2021	20/01/2022	100%	
14	3.1 Setup NodeRed	08/12/2021	31/12/2021	100%	Niklas Jaggy
15	3.2 Create DB Schema	07/12/2021	27/12/2021	100%	Niklas Jaggy
16	3.3 Data Management in DB	07/12/2021	27/12/2021	100%	Niklas Jaggy
17	3.4 Publishing Data as Service	29/12/2021	31/12/2021	100%	Niklas Jaggy
18	3.5 Create Metadata Catalog	11/01/2022	20/01/2022	100%	Simon Meyer,Niklas Jaggy
19	<b>4. Dashboard Implementation</b>	15/12/2021	25/01/2022	100%	
20	4.1 Dashboard Design	15/12/2021	25/01/2022	100%	Simon Meyer,Niklas Jaggy
21	4.2 Map Creation	31/12/2021	10/01/2022	100%	Simon Meyer
22	4.3 Service Integration	10/01/2022	18/01/2022	100%	Niklas Jaggy
23	<b>5. Project Documentation, Presentation &amp; Evaluation</b>	04/10/2021	31/01/2022	100%	
24	5.1 Create Project Presentation	17/01/2022	25/01/2022	100%	Niklas Jaggy,Simon Meyer
25	5.2 Create Project Report	20/12/2021	31/01/2022	100%	Simon Meyer,Niklas Jaggy
26	5.3 Project Documentation	04/10/2021	31/01/2022	100%	Simon Meyer,Niklas Jaggy
27					