TECHNICAL REPORT

European Food Safety Authority
Supporting Publications

APPROVED: 27 September 2022
doi:10.2903/sp.efsa.2022.EN-7639

Review of the Use of GIS in Public Health and Food Safety

European Food Safety Authority (EFSA),
Ana Francisca Jijón, Raquel Costa, Klara Nicova, Giandomenico Furnari

Abstract

The use of geographic information systems (GIS) is increasingly common in the scientific domains of public health and food safety. In this paper, we present the results of three analyses that we conducted to explore the application of GIS in these domains. The analysis includes a case study to test the feasibility of providing GIS services through an application developed with open-source tools. The first section presents the results from a scoping review looking at the current GIS practices in the public health and food safety domains as found in scientific literature. The second section shares the findings from a visual exploration of the use of GIS by EU agencies and supra-national bodies in the public health and food safety domains. The third section presents the current scenario of GIS used at the European Food Safety Agency (EFSA), which was tracked with an internal survey circulated among EFSA's scientific community. Alongside the three analyses, we look at current practices in terms of 1) fields in which GIS is used, 2) the infrastructure used for GIS, 3) GIS data, and 4) the methodology used for data collection, processing and analysis. The results show that GIS is used in multiple fields related to public health and food safety, such as epidemiology, plant health, animal health, and agriculture. We conclude that despite facing challenges regarding data quality issues, legal implications, and technical limitations, GIS has great potential given its relevance to the current environmental and health crises: such a holistic approach is in line with the One Health concept supported by EFSA and different European and international institutions. Moreover, GIS facilitates transition from control to prevention thanks to its capacity to develop spatial models that move from description to prediction. Finally, GIS facilitates communication, which guides better policy making.

Key words: GIS, public health, food safety, spatial data, spatial analysis

Requestor: EFSA

Question number: EFSA-Q-2022-00413

Correspondence: plants@efsa.europa.eu

Acknowledgements: EFSA wishes to thank the following for conducting this technical report: Ana Francisca Jijon (AJ) managed the analysis conducted for this report in collaboration with Klara Nicova (KN), Raquel Costa (RC) and Giandomenico Furnari (GF). AJ led the data collection, analysis and writing of results. KN supported data collection and processing and writing the results. RC supported data collection and performed the quantitative analysis. GF developed the application presented as a case study with the support of AJ.

EFSA wishes to thank all EFSA staff for the support provided to this technical report, in particular: Andrea Maiorano, Elisa Fasanelli, Luca Pasinato, Giulio Di Piazza, Jose Cortinas Abrahantes, Irene Munoz Guajardo, Irene Da Costa, Fulvio Barizzone, and all the EFSA staff from the scientific community who participated in the survey on GIS.

This technical report is a deliverable of the SEED project.

Suggested citation: EFSA (European Food Safety Authority), Jijón, A F, Costa R, Nicova K, iDATA, PLANTS, Furnari G, 2022. Review of the Use of GIS in Public Health and Food Safety. EFSA supporting publication 2022:EN-7639**.** 80 pp. doi:10.2903/sp.efsa.2022.EN-7639

ISSN: 2397-8325

Summary

In this report we present the results from different analyses that were conducted to present how GIS is used in terms of the technology, data, methodologies and infrastructure applied to conduct spatial analysis in the domains of public health and food safety. The analysis is complemented with a case study for which a GIS application was developed using open-source tools to provide self-served services to access and use spatial data.

GIS stands for Geographic Information Systems. It is a computer system used for collecting, managing, processing, and visualizing geographic data. GIS's strength is that it can combine different kinds of data (from attribute to spatial data, e.g., vegetation and virus outbreak) in one map. This feature enables illumination of spatial patterns that would be otherwise harder to recognize and allows to analyse and understand spatial patterns and relationships, for example the effects the environment has on human health.

This report is composed of three analyses. First is a scoping review that explored the current uses of GIS in the domains of public health and food safety in Europe and North America. Second is a visual exploration at agencies and supra-national bodies related to the domains of public health and food safety to identify how GIS is used in these domains by public agencies. Third, the report presented the current GIS scenario at EFSA which was tracked through an internal survey that aimed at identifying the activities and fields that use GIS. Finally, based on the findings of the three analyses, a GIS application was developed as a case study to test the feasibility of providing spatial self-served services using open-source solutions.

The results show that GIS is increasingly used in public health and food safety thanks to the rising awareness of the geographical and temporal dimensions of health. There are numerous fields where GIS is used by agencies. Among others, these are human epidemiology and public health, agriculture, plant and animal health, and environment. GIS is mostly used in these fields to 1) conduct analyses, 2) facilitate communication, and 3) support and inform policy makers.

The results highlight the potential of GIS to provide a holistic look at events thanks to its ability to join data across disciplines and geographic boundaries. Also, the findings show that in the domains of public health and food safety attention is turning from control to prevention, and GIS supports this transition as it provides the medium to conduct spatial statistical models that allow to move beyond description to prediction. GIS spatial analysis facilitates data exploration and interpretation. Furthermore, the results show the potential of GIS to communicate complex scientific findings in a straightforward and interactive way by showing results in maps or story maps.

Regarding data, the analysis identified a transition towards open data solutions, manifested in the use of public databases. Previous technical barriers to GIS are being overcome by the increasing number of user-friendly geoportals that allow to access and use spatial data. Also, new methods and technological advances allow continuous collection of data at increasingly finer temporal and spatial resolution, which however implies more complex models, and increases privacy concerns. In terms of infrastructure, the results show that there is a move towards web-based applications and the use of cloud computing, which allows the rise of a web GIS that enables data integration thanks to the internet, creating a network of data and overcoming data siloes. However, storage and sharing of data online poses legal issues regarding confidentiality. Results identified that geoportals and geodatabases are a common GIS infrastructure located in the web, whose advantages are facilitating data access and visualization, supporting data interoperability, allowing interconnection across disciplines and geographic boundaries, providing a centralized repository, and providing self-served services to conduct exploration and general geospatial analyses that help overcoming technical barriers.

Furthermore, the results identified the challenges faced by GIS. There are still no standardized practices to process data, nor harmonized models to analyse it, which hinders replicability of results and validity of outcome. Lastly, spatial data is still limited by data fragmentation and siloes, because of competing interests and variety of formats. There is an urgent need for centralized access to data. On the one hand, the growing number of portals and data access solutions supports data shareability. On the other hand, it complicates searching for data. Well-documented metadata is crucial for ensuring that data can be found and reused.

Moreover, this report explores the current use of GIS at EFSA. It recognizes that GIS is increasingly used at EFSA, particularly in the domain of biological monitoring. This section also identifies an external collaborative project between EFSA and other European bodies that uses GIS; the Information Platform for Chemical Monitoring (IPCheM) that has a spatial component. The results stress the importance of the development of GIS services and tools to support EFSA. Besides, EFSA aims to reach centralized geographic data management processes and practices that allow spatial data interoperability and models harmonization. This report includes a list of recommendations for future implementation of GIS at EFSA.

Lastly, based on the findings of the scoping review, the visual exploration, and the tracking of GIS use at EFSA, a GIS application was developed as a case study to test the feasibility of providing self-served GIS services to access and use spatial data using open-access tools. Several advantages have been identified regarding the implementation of a GIS application to offer spatial services. It has been relatively simple to develop a tool that visualises spatial data using open-source tools without depending on private software licenses. The application runs

a user-friendly interface that allows to easily (or automatically) select and plot data in an interactive, visually appealing and straightforward map. The application gives direct access to EFSA data, procuring a centralized flow of spatial data, while respecting its privacy policies. However, open-source libraries might also lead to disadvantages such as compatibility and maintenance issues.

The report presents the findings of each analysis, the scoping review, the visual exploration, track of GIS at EFSA, and the case study, in a separate section. Each section's findings are then brought together in a discussion where we present and compare the challenges faced by GIS and conclude that GIS is a powerful tool thanks to its capacity to combine data, enable spatial analysis, and support visualization.

Table of contents

1. Introduction

The use of geographic information systems (GIS) is increasingly present in the scientific fields of public health and food safety (Wang, 2020). GIS is a computer system used for collecting, managing, processing, and visualizing geographic data. GIS permits the integration of data, technology, people and processes, which enables combining spatial data, such as health (outbreak locations) and environmental data (EsriEUSummit, 2021). This helps to identify insightful spatial patterns and model the contagion and diffusion of specific diseases (Fatima et al., 2021).

Thanks to the advances in technology and methods to integrate, analyse, and visualize spatial data, attention in the current geospatial environment is turning from control to prevention (Kanankege et al., 2020). Furthermore, GIS plays an increasingly critical role in tackling the current climate crisis, furthering adaptation planning and implementation, as it allows the integration of data coming from different areas of study, such as the food sector, energy, built environment, and mobility (EsriEUSummit, 2021). In this way, GIS allows to respond to the challenge with a systemic change achieved through a holistic, transdisciplinary approach to health and sustainability, in line with the One Health concept supported by EFSA and different European and international institutions (Bronzwaer et al., 2021).

The use of GIS is increasing in EFSA. Therefore, there is a need for a more comprehensive view of the ongoing GIS activities and projects. Given the relevance of the topic, EFSA started actions to harmonize GIS data and its use to ensure better preparedness for future scientific assessment. Harmonized use of spatial data is a resonant theme throughout European institutions, including the European Commission with its INSPIRE Directive[1].

The main objective of the report is to view EFSA's GIS scenario into perspective with the current use of GIS in the domains of public health and food safety in terms of:

- Main fields where GIS is applied, whereby field is defined as a subfield inside the domains of public health and food safety such as spatial epidemiology or crops safety management.
- Infrastructure, referring to the type of technologies used to collect, manage and analyse data, differentiated by hardware2 and software3.
- GIS data, specifically aiming to identify the source, degree of accessibility, periodicity of data compilation, type and scale of data, as well as its ownership.
- Methodology for collection, processing, and analysis of spatial data.

The technical report builds on the following four activities:

- A scoping review looking at the current use of GIS in the public health and food safety domains as found in scientific literature.
- A visual exploration of public agencies' websites to identify any presence of GIS use.
- Tracking the current GIS projects and activities at EFSA through an internal survey.
- Case study demonstrating the feasibility of applying an open-source tool providing GIS services.

2. Data and Methodologies

2.1. Scoping review

The purpose of the scoping review was to identify the current GIS practices in terms of data, infrastructure, and methods used for public health and food safety academic studies in Europe and North America. Thus, the review aimed at answering the question:

How is GIS used for the study of public health and food safety in Europe and North America as published in the scientific literature during the last 5 years?

---

[1] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)

[2] Hardware: Tool or artificial body used for the collection, processing, storage and analysis of data.

[3] Software: Applications, data, and programs that run in a device used to collect, process, and analyse data.

A systematic search of scientific literature was conducted to retrieve the reviews from the following sources (Table 1):

Table 1: Data sources

Interface	Information Source
Web of Science - All databases	Web of Science Core Collection (WoS CC)
	BIOSIS Citation Index (BCI)
	Centre for Agriculture and Bioscience International (CAB Abstracts)
	Food Science and Technology Abstracts (FSTA)
	Medical Literature Analysis and Retrieval System Online (MEDLINE)

The search was limited to reviews in English published during the last five years to restrict the amount of literature to a manageable number. The search terms used to retrieve the reviews were the following: *Approach, Practice, Method, Process, Model, Scenario, GIS, Geographic Information System, Geospatial, Spatial data, Spatial information, Food Safety, Public Health, Epidemiology, Environment, Climate*. The final search strings are attached in Table 25, AppendixIII.

The first phase consisted of screening of abstracts and titles based on the inclusion criteria (Table 2) The second stage entailed reading the full text of the included articles, assessing them for relevance. Finally, data extraction was done using DistillerSR (Evidence Partners Incorporated; Ottawa, Canada).

Out of 205 retrieved articles, 26 reviews met the inclusion criteria and were therefore included in the analysis.

Table 2: Inclusion criteria for the scoping review

Field	Criteria
Study design	Studies designed to evaluate, assess, overview, and explore GIS practices.
Study characteristics	Assessment, reviews
Population/ Realm of interest	Public health, Food safety fields
Outcome of interest	Domain, infrastructure, data, and methodology
Language	English
Time	Last 5 years
Publication Type	Reviews
Geographic limitation	Europe and North America

2.2. Visual exploration

The main aim of the visual exploration was the identification of the use of GIS in public agencies and Supra National Bodies[4] from the domain of public health and food safety. In the end, 33 agencies' websites were included in the visual exploration (these include divisions from the same agency). The agencies were selected based on the inclusion criteria presented in Table 3.

Table 3: Inclusion criteria for the visual exploration

Field	Criteria
Material for exploration	Any evidence of use of GIS such as maps, story maps, interface providing GIS services
Source	Public agency from the public health and food safety domain
Outcome of interest	GIS products such as static, interactive and story maps within reports or published online. The existence of an interface with GIS services such as spatial data visualisation. Any material that evidences the use of GIS
Language	English
Time	Last 5 years
Publication Type	Website
Geographic limitation	Europe and North America

---

[4] For the continuation of the report, these will be referred to collectively as "agencies".

The European and US agencies and bodies that were searched are presented in Table 4. Besides, this section includes information retrieved from two conferences: 2020 GIS in EU virtual Summit and 2021 GIS in EU virtual summit .

Table 4: Agencies and supra national bodies selected

Region	Agency and Supra National Body
Europe	ECDC
	EEA
	EUROSTAT
	EC – DG
	JRC
USA	FDA
	CDC
	NCI
	USDA

ECDC (European Centre for Disease Prevention and Control), EEA (Environmental European Agency), EUROSTAT (European Statistical Office), EC (European Commission) – DG (Directorates-general departments), JRC (Joint Research Centre), FDA (Food and Drug Administration), CDC (Centers for Disease Control and Prevention), NCI (U.S. National Cancer Institute), and USDA (U.S. Department of Agriculture).

The visual exploration follows the same structure for data extraction as the scoping review. Hence, the same form in DistillerSR was used to document the visual exploration.

It is important to reiterate that the visual exploration was conducted informally. This means it was conducted within a structured framework but was constrained by the amount and accessibility of the resources available on the public agencies' websites. It is also important to note that the nature of this exercise was explorative and therefore might not form a comprehensive assessment of the GIS scenario in institutional agencies.

2.3. Tracking current situation at EFSA

An internal survey was circulated with the EFSA scientific community with the purpose of identifying the current EFSA GIS activities and projects, as well as the staff involved in these. Another main objective of the survey was assessing the level of harmonization and standardization of spatial data and models, to prepare for future implementation of GIS at EFSA. Besides, it aimed at collecting staff experiences and expectations for future GIS implementations.

The survey was divided in four sections: theoretical, technical, descriptive and the future scenario. The theoretical and descriptive section aimed at exploring the fields where GIS is used, the reason for applying it, and the legal implications of spatial data. The technical section inquired on the data flow, format of the spatial data, and tools used to conduct spatial analysis using GIS. The last section explored the future GIS scenario at EFSA. It aimed at capturing the problems encountered when using GIS and asked about the expectations for its future implementation.

The survey was distributed as a Microsoft Teams form through email. It circulated in two stages. First it was circulated in September 2021 with a sample group of 22 staff who were identified as GIS users through a snowball effect. The respondents were given two weeks to submit the survey. After having received the responses, it was decided to circulate the same survey for a second time addressing the whole EFSA scientific community in an attempt to reach a more representative sample of GIS scenario at EFSA. This second circulation took place in January 2022. During this round 12 additional responses were received out of which 2 were GIS users.

The responses to the survey are presented in an aggregated manner to protect respondents' anonymity. No further analysis was conducted on the obtained responses. Instead, these were observed as a descriptive snapshot of the current GIS situation at EFSA.

Results

3. Scoping review

A scoping review is a methodology used to collect as many insights as possible on a broad, or less reviewed subject. It is not limited to scientific articles and allows to gain a rather general picture of a subject without answering to a well-defined question. Given the explorative nature of this report, it was decided to conduct a scoping review to gain a general overview on how GIS is used in the domains of public health and food safety. Besides, a scoping review aims to achieve a better awareness of a topic, in this case, the use of GIS and its ongoing practices, and guides subsequent work.

3.1. Fields of GIS application

The reviews collected in this scoping review have been grouped by their subject of study. Three general fields have been identified. The fields identified are: 1) Public Health, with an emphasis on spatial epidemiology, 2) Ecology, and 3) Agriculture and Soil science in relation to public health. This section will highlight the application and justification for using GIS in these fields as described in the selected articles. These fields are by no means mutually exclusive. Indeed, they are all intertwined, which nicely demonstrates GIS ability to combine data across disciplines to apply a holistic approach to health and food safety. However, grouping reviews by subjects of study presents a good perspective to identify the use and advantages of GIS in different fields.

3.1.1. Public health

Several included papers discuss the application of GIS in the sector of public health. Public health, is defined as the science of preventing disease and promoting society health through organized efforts (WHO, 2022). The reviews show how geographic issues have become increasingly relevant to public health studies and policy, particularly through the lens of GIS (Wang, 2020). Hence, the selected reviews discuss health in relation to space, by giving attention to the environment's effects on health. Indeed, health is shaped by factors beyond genetics and clinical care such as sociodemographic and context conditions, so addressing environmental exposures through GIS will allow a better understanding of risk stratification of the population (Korycinski et al., 2018). For example, NASA's Earth Science area is moving towards the study of climate and environmental change and its impact on health (Bergquist and Manda, 2019).

Some of benefits that GIS brings to the study of public health are:

- allowing to assess the environment's effect on health across the time – space continuum (Chaix, 2018); Wang, 2020);
- helping to visualise data in ways that can reveal hidden relationships of health-space patterns (Bergquist and Manda, 2019; Fatima et al., 2021);
- revealing health disparities across space and demographic contexts (Korycinski et al., 2018; Card et al., 2019; Samuels-Kalow and Camargo, 2019; Wang, 2020);
- identifying how the environment influences access to health care and outcome of health interventions (Card et al., 2019; Wang, 2020);
- highlighting geographic heterogeneity of health (Wang, 2020).

Besides, a big advantage of GIS infrastructure is that it allows to combine attribute with geographic data across different disciplines through one common factor: the spatial component; thus, reducing data siloes (Delgado et al., 2019).

In public health, the rising awareness of the role of geography in the study of health has opened two novel research brands: neighbourhood's effect on health (effects of the neighbourhood's-built environment and community on an individual's health) and momentary analysis (the assessment of individuals' exposure to the spatial context across a time continuum, aimed at setting the basis for a life-segment epidemiology (Wilkins et al., 2017; Chaix, 2018; Wang, 2020). Time geography, an approach that looks at the relationship between space and time to explain the nature of events, is becoming more relevant in the study of public health. Time geography can also help to assess continuous exposure and predict health outcomes (Chaix, 2018).

3.1.1.1. Epidemiology

As a branch of public health, epidemiology addresses the relevance of the relationship of health and space. Epidemiology is the study of the causes of health outcomes and diseases in populations. It is the scientific study of the distribution (frequency, pattern) and determinants (causes, risk factors) of health-related states and events in specified populations (CDC, 2016). Epidemiology is also crucial for informing control and prevention capacity. Epidemiology is a broad subject, whose branches include, among others, human, plant, animal, and environmental epidemiology, the last dealing with the impact of environmental exposures on human health (Bloom, 2019). This section discusses the role of GIS in spatial epidemiology.

The geographical place and temporal dimension at which health events emerge and disperse have gained relevance (Cuadros et al., 2021). Spatial epidemiology is the study of health and disease over time and space (Fatima et al., 2021), and GIS enables incorporating the spatial dimension into the design and analysis of health across the continuum from prevention to treatment (Kirby et al., 2017). It has emerged as a novel approach for understanding health events and its association with the community and environment (Fatima et al., 2021). Spatial epidemiology also considers environmental factors such as land cover, water bodies, and air pollutants, and their effect on health (Jia et al., 2019a). Further, attention is turning from control to prevention thanks to the advancements of statistical geo-techniques (Kanankege et al., 2020). Indeed, GIS has been fundamental in guiding prevention measures during the COVID-19 pandemic and will continue being crucial beyond the COVID-19 crisis (Fatima et al., 2021).

According to Fatima et al. (2021), spatial analysis in epidemiology can be classified into three categories:

- Disease mapping or visualization, which describes geographic patterns of diseases.
- Exposure mapping, which explores the spatial covariance of potential influencing factors driving disease outcomes.
- Spatial epidemiological modelling, which aims at predicting health outcomes while adjusting for potential influencing factors and the spatial structure of the data.

Some of the benefits of spatial epidemiology are:

- identification of health patterns and drivers of disease distribution and transmission across time and space (Bergquist and Manda, 2019; Kanankege et al., 2020; Cuadros et al., 2021), and facilitation of interpretation of the results thanks to mapping and modelling (Bergquist and Manda, 2019; Malone et al., 2019);
- recognition of the fact that disease is unequally distributed across populations in geographical clusters given social, demographic and environmental differences (Jia et al., 2019a; Malone et al., 2019; Fatima et al., 2021; Cuadros et al., 2021);
- identification of vulnerable populations and indication of where interventions should be implemented (Cuadros et al., 2021);
- prediction of disease evolution and clustering in space over time (Kirby et al., 2017).

These benefits inform better policies on disease control and prevention (Malone et al., 2019).

Spatial epidemiology has been a successful approach to control infectious disease epidemics, and it has expanded to include the study of non-communicable diseases (Korycinski et al., 2018). Communicable diseases are transmissible diseases that result from infection. Several of the retrieved articles assess the use of GIS in the study of communicable diseases: COVID-19, HIV, and vector borne diseases such as malaria and leptospirosis (Dhewantara et al., 2019; Malone et al., 2019; Odhiambo et al., 2020; Fatima et al., 2021; Franch-Pardo et al., 2021). Spatial epidemiology proves essential in understanding the ecological and real time distributions of causative agents and their vectors (bacteria, parasites and viruses) at a high degree of accuracy (Bergquist and Manda, 2019). In cases of pandemics as the recent Covid-19, GIS can help to "monitor, evaluate situations, predict events, and inform policy decisions" (Franch-Pardo et al., 2021). Disease mapping has proven beneficial for detecting the spatial and temporal distributions of diseases, and for the exploration of the relationships between disease and socio-demographic factors (Fatima et al., 2021).

Behavioural, health related, and environmental risk factors are also linked to the increased risk of non-communicable diseases (Cuadros et al., 2021). The reviews assess the use of GIS to study diabetes, cancer, obesity and asthma, among others, and identify the risk factors in relationship to space (Jia et al., 2017); Korycinski et al., 2018; Card et al., 2019; Jia et al., 2019a,b; Sahar et al., 2019; Samuels-Kalow and Camargo, 2019; Cuadros et al., 2021).

Taking all the subtopics into consideration, GIS is used in human epidemiology to monitor population health and its relation to space.

3.1.2. Ecology

Two papers retrieved in this study show how GIS is used in the field of ecology. One of the subfields of ecology that relies heavily on the use of GIS is movement ecology (Seidel et al., 2018). It seeks to answer questions regarding where, how, and why animals move across different environments (Seidel et al., 2018). GIS aids that mission by providing methods for quantitative analysis of movement. In ecology, tracking the movement of animals is relevant because it influences spatial processes such as distribution of resources, disease transmission, and human–wildlife conflict (Seidel et al., 2018). Thus, the incorporation of movement ecology in epidemiological models enables more accurate predictions of spatio-temporal disease dynamics (Seidel et al., 2018).

Another purpose for the use of GIS is estimation of animal populations from remote sensing imagery using manual or automatized techniques (Hollings et al., 2018). GIS data on animal populations is essential for:

- Endangered species conservation
- Tracking invasive species spread
- Biosecurity
- Wildlife monitoring

Hence, GIS opens the possibility to study ecology in terms of animal populations, behaviour and movement, which contributes to the understanding and study of health in relation to emergent spatial processes and patterns.

3.1.3. Agriculture and soil science

Humanity is confronted with the big challenge of increasing agricultural production to achieve food security, while assuring the sustainability of food systems in face of climate change, addressing depletion of water resources, potential for increased erosion due to extreme weather events, and potential increase of the content of toxic elements in the soil due to anthropogenic activities and natural causes (Delgado et al., 2019).

Geographical data in soil science helps to:

- provide information on the state of health of the soil/sediment to policymakers or end-users (Agyeman et al., 2021);
- delineate the distribution of potentially toxic elements, their concentration, occurrence and possibly the source of pollution (Agyeman et al., 2021);
- identify temporal changes or development of soil health (Agyeman et al., 2021);
- support accuracy for sustainable and precision agriculture and environment (Agyeman et al., 2021);
- support the analysis of spatial and geo-temporal trends of ecosystems variables, e.g. climate dynamics, soil properties, variability, vegetation form distribution (Agyeman et al., 2021);
- model crop production and support crop surveillance, because crop types and their environmental hardiness vary regionally by latitude (Delgado et al., 2019);
- conduct climate and habitat suitability analysis, species distribution and niche modelling (Bergquist and Manda, 2019);
- evaluate and predict climate change effect on existing and established species (Bergquist and Manda, 2019);
- guide risk assessment and studies of invasive plants (Bergquist and Manda, 2019).

Furthermore, GIS allows connecting local data (e.g. from the farm) to the regional view of agriculture (e.g. the laboratory), which helps transitioning from a site-specific management focus to a notion of global sustainability (Delgado et al., 2019). Similarly, given the spatial nature of agriculture and remotely sensed data, GIS helps minimizing data siloes by providing spatial context around data (Bergquist and Manda, 2019).

To conclude, GIS is used in agriculture to enable sustainable, precise and efficient production and conservation. This is possible thanks to its modelling capacity and the availability of georeferenced climate and soil data.

3.2. Data

Spatial data is arguably more available than ever before. This analysis strives to capture the current state of GIS data. It explores the challenges, and limitations of geographical data, and presents some future needs and recommendations as found in the reviews.

3.2.1. Current state

3.2.1.1. Data Resolution and Scale

The findings of the scoping review show that on the one hand, the advances in data collection tools such as satellite imagery and sensors, have opened an unprecedented level of sophistication, creating data at higher time and spatial resolution and with more frequent coverage (Seidel et al., 2018; Hollings et al., 2018; Jia et al., 2019a,b; Malone et al., 2019; Odhiambo et al., 2020). This data is becoming increasingly accessible, and, in some cases, freely available (Jia et al., 2019a) which enables novel mapping and statistical techniques to be used to inform public health policy (Odhiambo et al., 2020).

The spatial unit of analysis is conditioned by the data availability and resolution. The findings of the scoping review show that data at a small scale (sub-national, national and regional) is mentioned more frequently across the assessed studies. However, several reviews also mentioned the use of larger scale (unit of analysis at city or point level; Figure 1). This higher prevalence of small scale over large scale unit of analysis is probably due to the legal and confidentiality limitations implied in high spatial resolution data, and to the fact that governmental databases often collect data at the subnational level, a scale that is used in epidemiological studies (Odhiambo et al., 2020).

Although the geographic unit is constrained by the available data resolution, Samuels-Kalow and Camargo (2019) state that it is usually recommended to use the smallest unit of analysis possible. However, Sahar et al. (2019) suggest that users should seek data at the geographic level that best aligns with the scale of the analysis. Nevertheless, another review adds that health data from predefined administrative or census units have limited value to the public that desires a comprehensive overview of a region, or to researchers who are interested in patterns at finer geographic scales (Wang, 2020). Jia et al. (2019a) sum it up by saying that studies using GIS data at a small scale, are limited by the spatial resolution of the data, while those using data at a large scale are constrained by the frequency of the data collection. Besides, considerations influencing choice of data source may include cost, accessibility, geographic scope, age, and quality of data (Wilkins et al., 2017).

Regardless of the choice of the scale, careful consideration should be taken when integrating data at different geographic levels to avoid unnecessary assumptions, such as uniform distribution across a region (Sahar et al., 2019). Furthermore, it is important for researchers to keep in mind that choosing a different unit of analysis can lead to different conclusions from the geographic analysis, known as the modifiable aerial unit problem (MAUP; Kirby et al., 2017; Korycinski et al., 2018; Bergquist and Manda, 2019; Samuels-Kalow and Camargo, 2019).

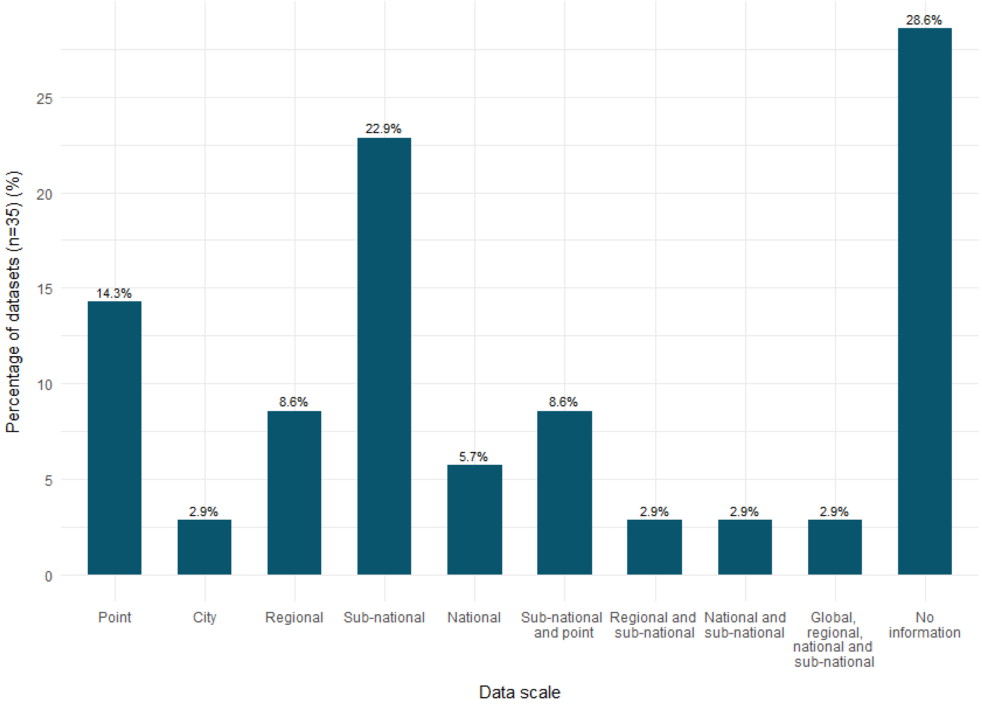


Figure 1: Data scale of the datasets mentioned in the reviews

When it comes to the periodicity of data collection, the scoping review identified that 11.4% of datasets mentioned the use of continuous data (i.e. continuous collection of spatial and time data thanks to the advances of mobile devices and apps, citizen science, crowdsourcing, wearable sensors, etc.; Figure 2). Continuous data was mentioned in studies on environment, mobility, geographic momentary assessment, soil science, and human epidemiology (Chaix, 2018; Bergquist and Manda, 2019; Odhiambo et al., 2020; Wang, 2020; Agyeman et al., 2021).

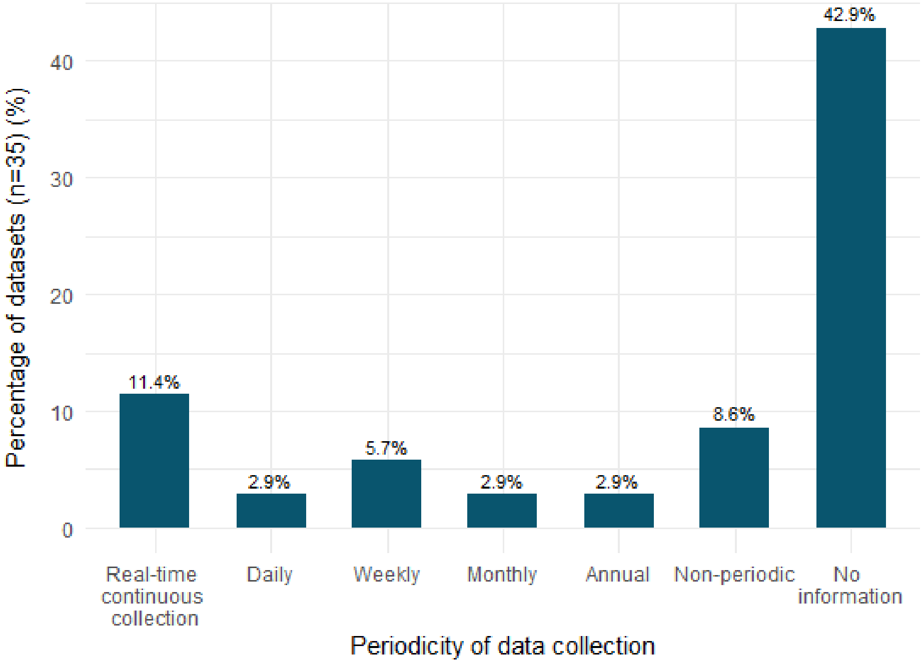


Figure 2: Periodicity of data collection of the datasets mentioned in the reviews

Please note that due to the nature of the data collected from reviews information regarding data scale and collection is limited, with almost 29% of no information for data scale and 43% for data collection periodicity.

3.2.1.2. Data Source

The scoping review findings show that public datasets are more common than private ones (37.2% vs. 28.5%; Figure 3). The higher reliance on public datasets might be related to the current interest in open data, data shareability, interoperability, and results reproducibility within the scientific community (Seidel et al., 2018; Hollings et al., 2018; Card et al., 2019; Odhiambo et al., 2020). However, just because a dataset is private doesn't have to mean its data is not accessible and vice versa. While 69.2% of the public datasets provide open access to the data, 40 % of the private datasets provide open data (Figure 3). It is important to consider that using different data sources in an analysis can lead to different conclusions, given differences in metrics definitions, and demographics data (Wilkins et al., 2017).

Note that 35% of data sources provide no information because the reviews did not specify the data source or accessibility.

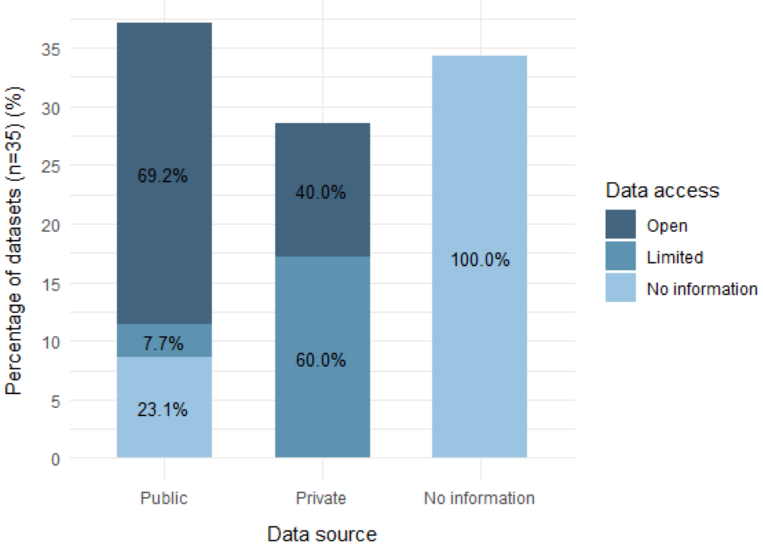


Figure 3: Data access and source characteristics of the datasets mentioned in reviews

3.2.1.3. Data Type

The scoping review has identified the use of different data types. GIS data can be either presented as vector, raster or spatio-temporal (Bergquist and Manda, 2019). GIS allows to combine attribute with spatial data using location as the common factor (Delgado et al., 2019).

- **Attribute data:** Nonspatial information about a geographic feature in a GIS, usually stored in a table and linked to the feature by a unique identifier (ESRI).
- **Vector:** composed of points, lines and polygons and a combination of attribute and spatial data.
- **Grid:** composed of pixels (often referred to as raster); often used for geographical base maps as well as continuous phenomena such as land cover, digital elevation models and climatological information (Bergquist and Manda, 2019).
- **Spatio-temporal:** Vector or raster containing spatio-temporal information. Locations in space may be recorded by coordinates representing longitude, latitude and elevation at a given time.

The scoping review findings show that attribute, spatial, grid, and temporal data are often used in combination. The data types identified in the scoping review refer to the data that was used by the studies assessed in the selected reviews. The most common data type in GIS analysis is attribute in combination with spatial information. Interestingly, grid data seems to be used independently according to the findings from the scoping review (Figure 4).

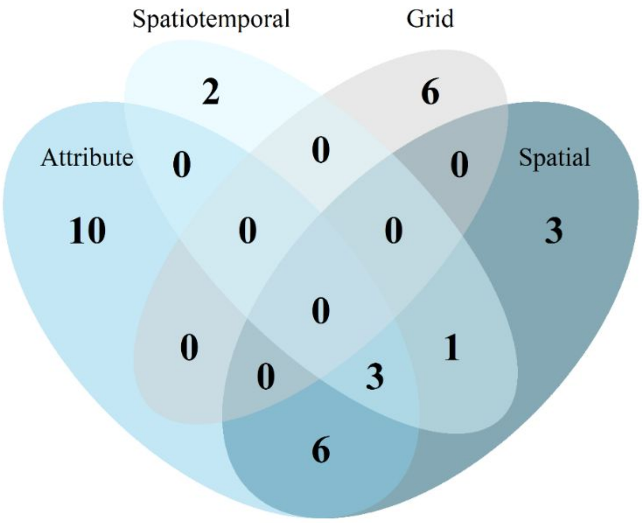


Figure 4: Combinations of data type use from the datasets mentioned in reviews

3.2.2. Challenges and limitations of geographic data

The following section presents the challenges and limitations that have been identified in the scoping review. On the one side, data collected at higher space and temporal resolutions brings advantages due to the detail of information. On the other hand, the immense amounts of data collected and accessible at increasingly higher resolution also pose challenges regarding the methods to use to store, and analyse GIS data. Also, higher spatio-temporal resolution implies legal limitations regarding privacy and confidentiality.

3.2.2.1. Methods

As more data is becoming available in higher spatial and temporal resolution, there is the need for more complex and effective modelling methods to maximize gained insights and minimize uncertainties (Seidel et al., 2018; Bergquist and Manda, 2019; Odhiambo et al., 2020; Agyeman et al., 2021). Accordingly, there is a need for novel statistical approaches and effective prediction methods that could unlock the value of rich geospatial data (Korycinski et al., 2018; Agyeman et al., 2021).

3.2.2.2. Legal limitations: privacy and confidentiality

As the spatial resolution of the data gets finer, the policies related to privacy and confidentiality increase, making “geoprivacy” an urgent issue in all stages of conducting a study of health (Wang, 2020). There are political, legal and privacy issues with sharing data, especially when human subjects are involved (Seidel et al., 2018), and when data scale is at the level of an individual (Kirby et al., 2017; Korycinski et al., 2018). Many public datasets are modified to protect confidentiality by applying:

- statistical methods, such as data swapping (switching values between records, referred to as geomasking when using geographic data), cell suppression (excluding data), top coding (reporting values as "above" a certain threshold), and "rounding" (Sahar et al. (2019)
- Data aggregation

Despite the use of data-masking techniques to protect confidentiality, Sahar et al., (2019) mentions that it is preferable to use dot maps with exact addresses only for inhouse exploratory analyses and avoid publishing them. Hence, although more and more data is collected at finer spatial scale, privacy and confidentiality issues limit the publication of such data.

3.2.2.3. Lack of standardization

Despite the existing attempts to implement a common geographic system format and language like the Open Geospatial Consortium's Geography Markup Language (GML; (OGC, 2022b), the lack of standardized approaches to data selection and sharing remains a limitation to explore the potential of spatial data (Fatima et al., 2021). Accordingly, inconsistent measuring standards and collection methods among different data providers make it hard to combine data from different sources (Jia et al., 2017). Jia et al. (2017) add that particularly commercial databases suffer from a critical weakness due to different data providers' inconsistent measuring standards or collection methods.

3.2.2.4. Data fragmentation and data siloes

Another challenge faced by spatial data is fragmentation and creation of siloes. Continuous, big data collected from satellites, sensors or GPS sensors is still fragmented due to competing interests and variety of data formats. Also, as the number of available geodatabases increases, so varies data volume and velocity. This has led to siloed data management infrastructure (Delgado et al., 2019). Nevertheless, according to Delgado et al., (2019), a web GIS platform[5] provides a framework for reducing past siloes across the public sector agencies, but also between the public sector and agricultural industry (Delgado et al., 2019). Besides, several initiatives exist in order to overcome data fragmentation and siloes by creating a centralized digital platform to access thematic information such as the European Strategy for Data[6], the aligned INSPIRE Directive[7] and the Destination Earth (DestinE) initiative[8].

3.2.2.5. Data uncertainties and validity of outcome

The current generation of large and continuous amounts of data leads to problems like uncertainty and inaccuracies in the generated data (Bergquist and Manda, 2019). Also, spatial uncertainty in data and analysis arises due to measurement errors, generalization of spatial features and "geoprivacy" issues from crowdsourced data (Wang, 2020). In recent Covid-19 studies, data quality and ecological fallacy was the main limitation to the use of spatial techniques (Fatima et al., 2021).

Regarding data coming from sensors, Bergquist and Manda (2019) state that "users should select data carefully as time series observation are usually affected by the density of measurement and non-homogeneity caused by non-climate factors". Also, data sources and acquisition style vary, which leads to non-reproducible findings (Fatima et al., 2021). Further, the validity of the outcome depends on the data type and spatial scale at which the analysis was performed (Dhewantara et al., 2019).

3.2.3. Needs and future recommendations

There is currently a strong call for:

---

[5] Web GIS platform is the infrastructure built by web geoportals and web geo data hubs interconnected through the internet

[6] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A European strategy for data

[7] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)

[8] Destination Earth (DestinE), initiative of the European Commission

- Open, centralized and easily accessible data (Dietrich et al., 2018; Seidel et al., 2018; Card et al., 2019)
- Reduction of cost of research
- Reproducibility of results
- Shareability of data and results
- Data interconnectivity to avoid data fragmentation
- Interoperability

This could be achieved with a web platform that would warrant:

- Data availability, interoperability and quality of data issued from different sources (Dietrich et al., 2018);
- Collection of rapid and accurate data at a fine-scaled geographic resolution (Bergquist and Manda, 2019);
- Facilitated application of the most advanced spatial approaches;
- Analysis at finer spatial level;
- Orienting of communicable diseases control and elimination efforts.

3.3. Infrastructure

This report understands infrastructure as the technical architecture of tools used to collect, manage, share, publish and analyse data. Infrastructure consists of hardware and software. Both are used in all stages of GIS studies, from data collection and processing to data analysis. Due to the limited scope of this review, the report mainly concentrates on hardware used for data collection such as sensors, satellites and drones, while for software it concentrates on the applications and programs used for data retrieval, processing and analysis.

Advances in technology have enabled the collection of data at finer spatial and temporal resolution. This possible leads to creation of large amounts of data which demand high hardware storage and computational capacities (Delgado et al., 2019), as well as software capable of handling big data (Hollings et al., 2018). There is also a trade-off between data resolution and cost (Jia et al., 2019a). Therefore, the choice of infrastructure depends on the research question as well as on the resources available.

The following sections present a non-exhaustive overview of hardware and software tools, which have been identified in the literature. However, it does not attempt to recommend the most suitable tools, because those always depend on the context.

3.3.1. Hardware

This report defines hardware as physical devices used for the collection, processing, storage and analysis of data (TechTerms, 2022). It mainly focuses on the devices used for data collection due to the nature of the information retrieved from the reviews.

3.3.1.1. Remote sensing

Remote sensing (RS) refers to the methodology of data collection at distance, usually from a satellite or an aircraft (Dietrich et al., 2018). The satellites and aircrafts carry sensors which detect electromagnetic radiation signal (Hollings et al., 2018). These signals can be further processed by algorithms to derive various parameters of interest such as temperature, altitude or vegetation indices (Dietrich et al., 2018).

The main advantages of remote sensing according to Hollings et al. (2018) are:

- large spatial coverage;
- faster data procurement with reduced costs;
- feasibility of regular revisits of sites;
- fewer requirements for human intervention;
- unobtrusiveness of data collection;
- possibility of data collection in places where in-situ sampling is unfeasible (in remote localities or due to social conditions like lock-downs; Franch-Pardo et al., 2021).

RS has been utilized in public health in three main ways:

- to identify associations between diseases and remotely sensed parameters
- to model and predict disease spread
- to establish early warning systems and carry out real time monitoring; (Dietrich et al., 2018; Delgado et al., 2019; Dhewantara et al., 2019; Jia et al., 2019b; Malone et al., 2019; Odhiambo et al., 2020)

Real-time monitoring is an emergent theme in the GIS community. It not only applies to data collection but also to the response to the data (Seidel et al., 2018; Malone et al., 2019; Franch-Pardo et al., 2021). The development of near real-time surveillance systems based on GIS, global positioning systems and RS facilitates the establishment of accurate, up-to-date early warning systems (Samuels-Kalow and Camargo, 2019; Malone et al., 2019).

However, even with these advances in RS, ground-truth data remain important, especially when features of interest cannot be easily distinguished from earth observation data (Jia et al., 2019a). Furthermore, the requirements for technical capabilities and software needed for an analysis of RS data are still limiting the uptake of the technology (Hollings et al., 2018).

It is important to note that different data collection technologies can be used together in a complementary manner. An approach combining the use of RS, Global Navigation Satellite Systems (GNSS) and GIS technologies can be used to study both communicable and non-communicable diseases (Dietrich et al., 2018; Jia et al., 2019a,b). For example, studies of non-communicable diseases have used wearable RS sensors to monitor individual's exposure to environmental factors across a time-space continuum with the aim to analyse their impact on health (Chaix, 2018, Bergquist and Manda, 2019, Fatima et al., 2021). This methodology enables studying individual lifestyles, behaviours and health outcomes at lower cost to researchers compared to traditional in-situ sampling (Jia et al., 2019b).

3.3.1.2. Hardware used for remote sensing

We had identified seven main types of hardware used for the collection of spatial data (Figure 5). These are global navigation satellite systems (GNSS), light aircrafts, radiometers, satellites, sensors, smartphones, and unmanned aerial vehicles (UAVs, also known as drones).

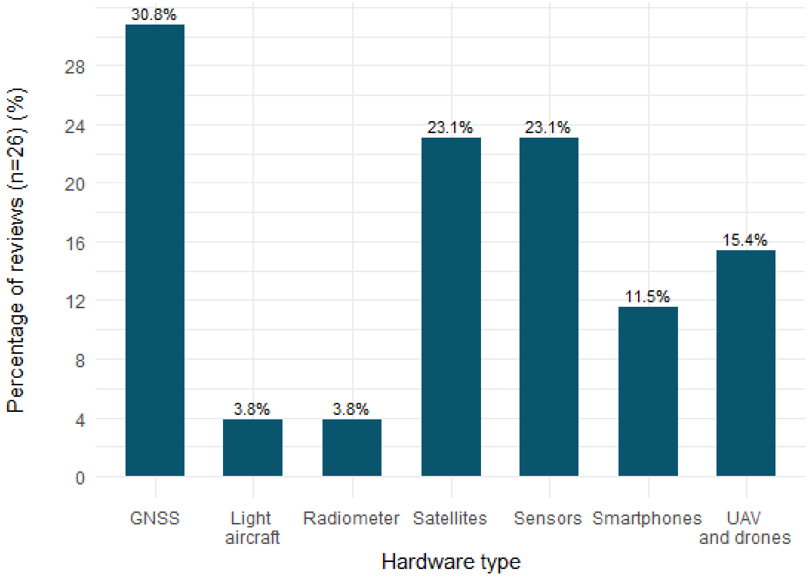


Figure 5: Percentage of reviews mentioning each type of hardware

Figure 5: Percentage of reviews mentioning each type of hardware.

Light aircrafts and UAVs

The main difference between a light aircraft and an unmanned aerial vehicle (UAV, sometimes also called a drone) is that the latter is operated remotely without any human pilot on board. Both technologies share the advantages of high spatial resolution and flexibility (Hollings et al., 2018). Their flights can be timed for specific events and they allow a relatively fast uptake of new developments in sensors and other technology (Hollings et al., 2018).

However, imagery taken from light aircraft is restricted to areas within its reach which tends to be limited by access to runways or fuel limits (Hollings et al., 2018). These restrictions are even more pronounced for UAVs which usually have a short line of sight from operator and require several passes to cover the same area as a light aircraft (Hollings et al., 2018). Furthermore, small UAVs can only carry smaller, often lower quality sensors which limits the spatial resolution they can achieve (Hollings et al., 2018). Light aircrafts can take on many sizes from quad-copters to full sized airplanes, although an increase in size goes in hand with increased costs and technical skills required for operation (Hollings et al., 2018).

Notwithstanding these limitations, UAVs have been used extensively to collect different types of environmental data of public health relevance (Jia et al., 2019a), as well as to study land use change (Franch-Pardo et al., 2021), or estimate animal populations (Hollings et al., 2018).

GNSS + Earth observation satellites

Satellites are artificial bodies placed in orbit around the earth, moon or another planet in order to collect information. They can operate by themselves or be a part of satellite constellation (NASA, 2014). Satellites are vital for data collection because they can monitor the entire globe within short time periods (Bergquist and Manda, 2019). Satellites can be divided into many subcategories; in this report we will separately discuss Global Navigation Satellite Systems (GNSS) and Earth observation satellites (EO). While EO are used for collecting information about the Earth from the orbit, usually in the form of aerial imagery, GNSS provide geo-spatial positioning information. They are discussed separately, because each of them has different uses, benefits, and challenges.

Earth observation satellites

Earth observation satellites are devices used for collecting information about the Earth from orbit, usually in the form of aerial imagery. Some satellites can collect imagery from around the globe on a daily basis by which they are narrowing the gap between satellite images and aerial photography (Jia et al., 2019a). The latest developments offer imagery with sub-meter resolution, which allows to construct risk assessments on the level of a community or even a household (Bergquist and Manda, 2019). Some satellites are capable of sharing data in real time (Malone et al., 2019) which makes it possible to conduct momentary analysis. Furthermore, many satellites provide free access to the data they collected, for example Terra and Landsat data are accessible to the public for free (Jia et al., 2019a). In addition to government operated satellites, there is a growing number of privately launched ones. Indeed, their number is expected to double by 2027 (Bergquist and Manda, 2019).

Even though satellite imagery has many advantages, it is also subjected to some limitations. The purchase of satellite data is sometimes linked to minimal order requirements which can considerably increase the cost, especially for smaller research subjects (Hollings et al., 2018). Furthermore, there might be gaps in satellite data due to financial or technological constraints, for example caused by cloud coverage. It might be possible to solve these limitations by complementing the missing data with aerial imagery taken from UAVs or light aircrafts, because their flights can be planned specifically according to the data needs (Jia et al., 2019a). While satellite data are ideal for large and remote areas where spatial resolution does not need to exceed 0.5 m, aerial imagery from UAVs or aircrafts is better suited for small scale applications where a higher resolution is desirable (Hollings et al., 2018).

Global Navigation Satellite Systems

Global Navigation Satellite Systems (GNSS) are a constellation of satellites providing signals from space that transmit positioning and timing data to GNSS receivers. The receivers then use this data to determine location (EUSPA, 2022). The most well-known examples of GNSS are American GPS, European Galileo, Russian GLONASS, Indian IRNSS, Japanese QZSS and Chinese BeiDou (Dietrich et al., 2018). GNSS data can be obtained from various sources such as smartphone location data, the internet and social networks (such as Google, Baidu, Facebook, Twitter geotagged tweets, Weibo, Quingbo), as well as from road, air and sea traffic monitoring (Franch-Pardo et al., 2021).

The main advantage of GNSS is global coverage and precision, which ranges from a few meters to some centimetres (Dietrich et al., 2018). Such continuous data enables the development of new research protocols that can track movement over time (Samuels-Kalow and Camargo, 2019).

GNSS has been successfully used for the study of both communicable and non-communicable diseases (Dietrich et al., 2018), by enabling to:

- determine individuals' locations, movements, modes of transportation, and activities (Jia et al., 2019a);
- measure individuals' exposure to risk factors for diseases (Samuels-Kalow and Camargo, 2019);
- link identified infection cases with patients' travel histories (Franch-Pardo et al., 2021);
- to measure levels of physical activity (in obesity research; Jia et al., 2019b).

- The requirements for the successful use of GNSSs are stable and easily accessible signals, as well as procedures preventing power failure (Dietrich et al., 2018).

Sensors

Sensors are here defined as devices that measure a physical quantity and convert it into a signal which can be read by an observer or by an instrument. Various sensors, including those incorporated into mobile phones, are particularly useful for studies exploring dynamics at a micro-temporal scale (Chaix, 2018). They allow researchers to study immediate determinants of health in behavioural contexts (Chaix, 2018). Various sensors have been utilized for the study of humans, animals as well as for precision agriculture (Chaix, 2018; Seidel et al., 2018; Delgado et al., 2019). Examples of sensors used in the field of food safety and public health include GPS tags, accelerometers and proximity sensors.

GPS tags[9] are useful for obtaining high-resolution spatio-temporal location data (Chaix, 2018). They are becoming more affordable which enables their use on a larger scale (Jia et al., 2019b). They can be further paired with accelerometers, proximity tags, radiometers and other sensors to produce more accurate data (Jia et al., 2019b). GPS tags are utilized by scientists studying humans as well as animals (Dietrich et al., 2018; Seidel et al., 2018; Card et al., 2019; Jia et al., 2019b). For example, movement ecologists use GPS tags to follow locations of samples of animal populations to gain insights on the effectiveness of management strategies for species conservation or mitigation of wildlife conflicts (Seidel et al., 2018). Proximity tags have also been employed by movement ecologists to advance the study of interactions between animals which has important implications for understanding disease transmission and competition for resources (Seidel et al., 2018).

The main challenges in the use of sensors are predicted to be their miniaturization, cost reduction, and increase in autonomy of the devices both regarding their battery-life and memory size (Chaix, 2018). Furthermore, there is a need to increase interoperability of the devices since presently they require many different cables, memory cards and software for accessing the data (Chaix, 2018). This issue could be bypassed by real-time transmission of data to servers; however, this may not be always feasible, and it could induce high costs (Chaix, 2018).

Smartphones

Smartphones with integrated sensors are a potential valuable source of information for studies in humans. They are programmable, have powerful processors, large amounts of memory and can be linked to external devices through wireless connections (Chaix, 2018). However, battery life remains an issue when several functionalities are used at the same time (Chaix, 2018). A combination of location data from GPS with various sensors embedded in smartphones has led to the development of many health-tracking applications for smartphones (Jia et al., 2019b). Such mobile health and mapping applications have many potential uses, for example one has been implemented to improve asthma self-management (Samuels-Kalow and Camargo, 2019).

3.3.2. Software

This report understands software as applications and programs that run in a computer or device and are used to retrieve, process, and analyse data (TechTerms, 2022). Five types of software have been identified in the reviews: geodatabases[10], geoportals[11], specialised GIS software, image processing software, and specialized computer programs or programming languages used for analysis in statistics or data sciences, hereby referred to as statistical software (Figure 6).

---

[9] Navigational devices using GPS to determine users' location

[10] Spatial database is a database that is designed to store, query, and manipulate geographic information and spatial data. For the scope of this research, we assume that geo databases offer discovery services. Download services are provided on a case by case basis.

[11] A geoportal is a type of web portal usually provided by a public organization, that is used to access geographic information and associated geographic services (display, editing, analysis, etc.) via the Internet. Geoportal users are usually not given complete access to the software suite but can access some services. Here we also consider Virtual Globes as a type of geoportal. For the scope of this research, we assume that geoportals offer visualization and discovery services. Download services are provided on a case by case basis.

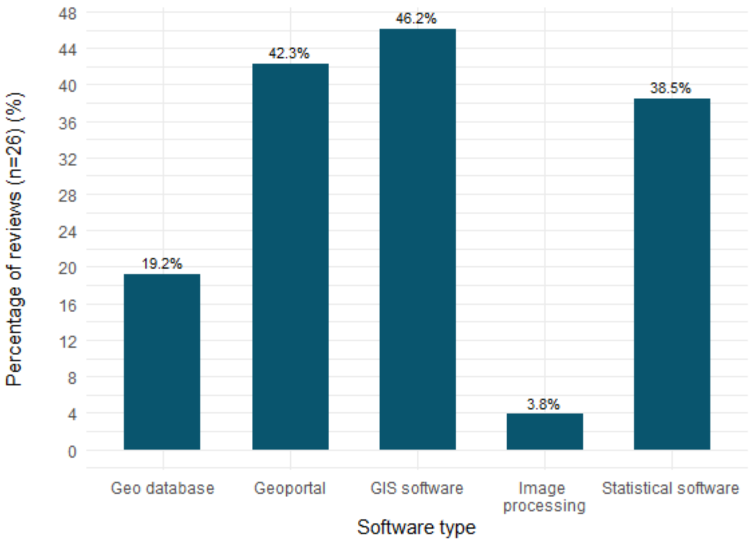


Figure 6: Percentage of reviews mentioning each type of software

Although the previously mentioned software can be complementary to each other, in this report we aimed at individually capturing which are the most commonly mentioned. The most mentioned software type is GIS software (mentioned in 46.2% of reviews), closely followed by geoportals (42.3%) and statistical software (38.5%). Among the reviews, there is a slightly higher mention of GIS software (e.g. ArcGIS and QGIS), and specialized computer programs or programming languages such as R that support spatial analysis (Table 5). At the same time, several specialized software solutions developed for different fields and types of analysis (e.g. Maxent, SATScan, OpenBUGS) can be used. Furthermore, there is still an untapped potential in using software developed for different disciplines such as biomedicine and engineering (Hollings et al., 2018).

Table 5: Most often mentioned software

Software name	Type	Distribution[12]	Licensing/Access	Number of reviews
ArcGIS	GIS software	Hybrid	Proprietary/cost	9
R	Programming Language	Hybrid	Open source/free	8
SaTScan	GIS software	Local	Proprietary/free	5
QGIS	GIS software	Local	Open source/free	4
GeoDa	GIS software	Local	Open source/free	3
Google Earth	Geoportal	Hybrid	Proprietary/cost and free	3
Google Maps	Geoportal	Distributed	Proprietary/cost and free	3
Python	Programming Language	Hybrid	Open source/free	3

3.3.2.1. Geoportals

A geoportal is here defined as a type of web portal, usually provided by a public organization, that is used to access geographic information and associated geographic services (such as display, editing, and analysis). Its users are usually not given complete access to the software suite but can only access certain services. The results show that geoportals usually offer visualization and discovery services, that means they allow to search and select thematic

---

[12] Distribution is here considered a generic identification of the process of delivering software to the end user rather than a specification of the software architecture. It can be local (application operates natively on devices in the local environment), distributed (application is accessed through the web), or hybrid (Location of the application is accessed through the web as also through the local application).

data to visualize it on a map. On the other hand, we specify which geoportals provide download services. Out of the 19 geoportals identified in the reviews, 79% enable users to download their data (Figure 7).

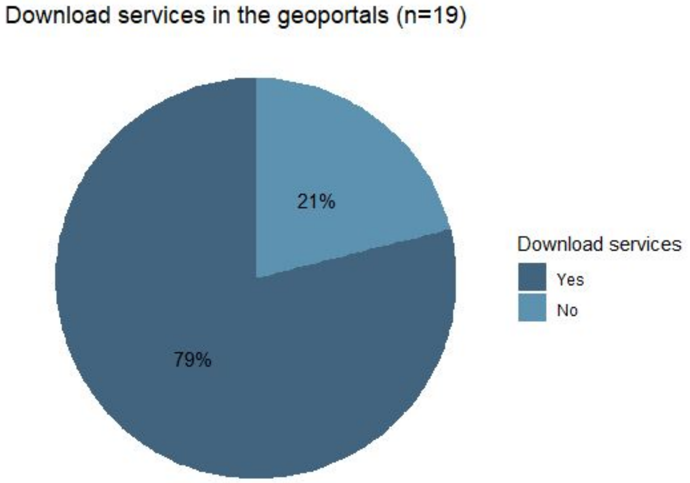


Figure 7: Percentage of the geoportals, mentioned in reviews, providing download services

A specific category of geoportals offering web-based map services are virtual globes. Virtual globes are interactive, user friendly platforms addressed at non-programmers (Broomhead et al., 2019). Two examples of such platforms are Google Earth Engine and Healthy futures atlas (Malone et al., 2019). Thanks to their ease of use, virtual globes have a strong potential to ease data sharing and dissemination of scientific research results (Bergquist and Manda, 2019; Jia et al., 2019a). The Google Earth Engine allows users to access and process satellite data using Google's cloud computing resources (Jia et al., 2019a). It incorporates archives of RS data which can be readily analysed even by researchers outside the remote sensing field (Jia et al., 2019a).

Many geoportals also enable users to create interactive maps. Examples of such platforms are the CDC State Cancer Profiles (Sahar et al., 2019), the CDC U.S. Cancer Statistics Data Visualization tool (Sahar et al., 2019), HealthMap (Kirby et al., 2017) or Darthmouth Atlas (Kirby et al., 2017). A platform called Global Disaster Alert and Coordinating System attempts to provide a comprehensive solution for early response to disasters including but not limited to the creation of disaster maps (Dietrich et al., 2018). Global Earth Information System of Systems is a platform which provides access to earth-observation data from different sources, thus increasing data availability (Dietrich et al., 2018). To see more examples of geoportals and their benefits, see section 4 *Visual Exploration*.

3.3.2.2. Geodatabases

In this remit geodatabases are defined as databases designed to store, query, and manipulate geographic information and spatial data. Our results show that geodatabases offer discovery services, i.e. allow users to retrieve spatial datasets based on metadata (InspireDirective, 2011). Most (93%) of the databases mentioned in the reviews also provided download services (Figure 8).

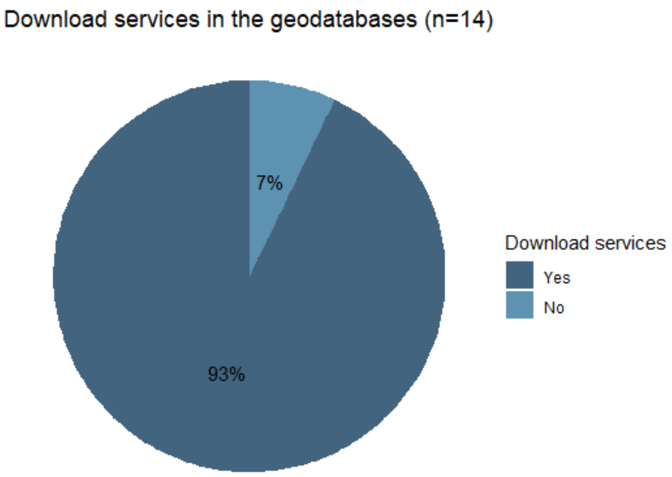


Figure 8: Percentage of the geodatabases, mentioned in reviews, providing download services

Several geodatabases have been identified in the scoping review. Food and Agricultural Organization (FAO) GeoNetwork provides publicly available GIS data from many fields including climate, agriculture, society and farming (Dhewantara et al., 2019). Movebank stores archives of animal tracking data and allows interactive mapping of animal movement (Seidel et al., 2018) (Movebank) (Movebank). Finally, Humanitarian Data Exchange is an open platform for data sharing in the humanitarian context (Dietrich et al., 2018).

3.3.2.3. Specialised GIS software

Specialised GIS software is designed to store, retrieve, manage, display, and analyse all types of geographic and spatial data. It is the most often mentioned software type, being mentioned in 46.2% of the reviews (Figure 6).

Popular uses of GIS software include geocoding (Kirby et al., 2017; Samuels-Kalow and Camargo, 2019; Sahar et al., 2019), creation of interactive maps (Kirby et al., 2017; Sahar et al., 2019; Franch-Pardo et al., 2021) or detection and visualisation of clusters (Kirby et al., 2017; Dhewantara et al., 2019; Sahar et al., 2019). ArcGIS is one of the most popular GIS software used in the field of public health and food safety (Table 5). Other popular options include SaTScan, QGIS and GeoDa (Table 5).

3.3.2.4. Statistical software

For the purpose of this report, we define statistical software as specialized computer programs or programming languages used for analysis in statistics or data sciences. As such, we discuss Python (a programming language), and R (an environment for statistical computing) under the umbrella of statistical software. The availability of programming languages and specialized computer programs allows complex spatial analysis to be conducted in an easier and faster way (Sahar et al., 2019). For example, Bayesian statistical methods are gaining popularity thanks to the advances in computational power and availability of software solutions (Bergquist and Manda, 2019). Most frequently mentioned examples of programming languages that support statistical analysis are R (31% of reviews) and Python (12%). Part of their popularity might lie in the availability of many packages (such as R-INLA or CARBayes for R, OSMnx for Python) developed for the purpose of conducting spatial analyses (Seidel et al., 2018; Broomhead et al., 2019; Sahar et al., 2019; Kanankege et al., 2020; Franch-Pardo et al., 2021). Packages for spatial analysis also exist in SAS (Sahar et al., 2019).

Finally, different software products can be used in a complimentary fashion. As Fatima et al. (2021) noted, most studies included in their scoping review used a combination of different statistical software products to conduct their analyses.

3.3.2.5. Image processing and classification software

Image processing and classification software are tools designed to manipulate digital images to enhance or extract information. Image classification software assigns objects in pictures into one or more categories based on previously defined properties (mvtec). Image processing and classification software provide a toolbox from transforming remote sensing imagery into useful map data, such as land use classes. Another use of image classification is the recognition of objects or animals (Hollings et al., 2018). Machine learning algorithms, both supervised and unsupervised can be used, often leading to improved classification results compared to manual classification (Hollings et al., 2018). Fields using this approach are ecology, agriculture, and also human epidemiology (Hollings et al., 2018; Delgado et al., 2019; Franch-Pardo et al., 2021). Some specialised GIS software such as QGIS and ArcGIS provide image processing and classification, which makes them the most utilized software for identification of animals from aerial imagery (Hollings et al., 2018).

3.3.2.6. Software distribution

Software distribution is defined in this remit as the process of delivering software to the end user. It can be either local, which means that the software needs to be installed and stored on a user's device, or distributed, meaning that the program is hosted on the internet or in a server, and provides services through a network connection. Some software, such as ArcGIS, combine both types of distribution in what is called a hybrid distribution. The scoping review shows that GIS applications, specially geoportals, geodatabases, and GIS applications are moving towards a distributed (web-based) distribution (Figure 9).

Web-based distribution is mostly utilized for geodatabases and geoportals, which were located on the web in more than 85% of the cases (Figure 9). On the other hand, specialised GIS software, statistical analysis software and image processing software are still mostly local. Hybrid distribution is mainly utilized for statistical analysis and image processing software (Figure 9).

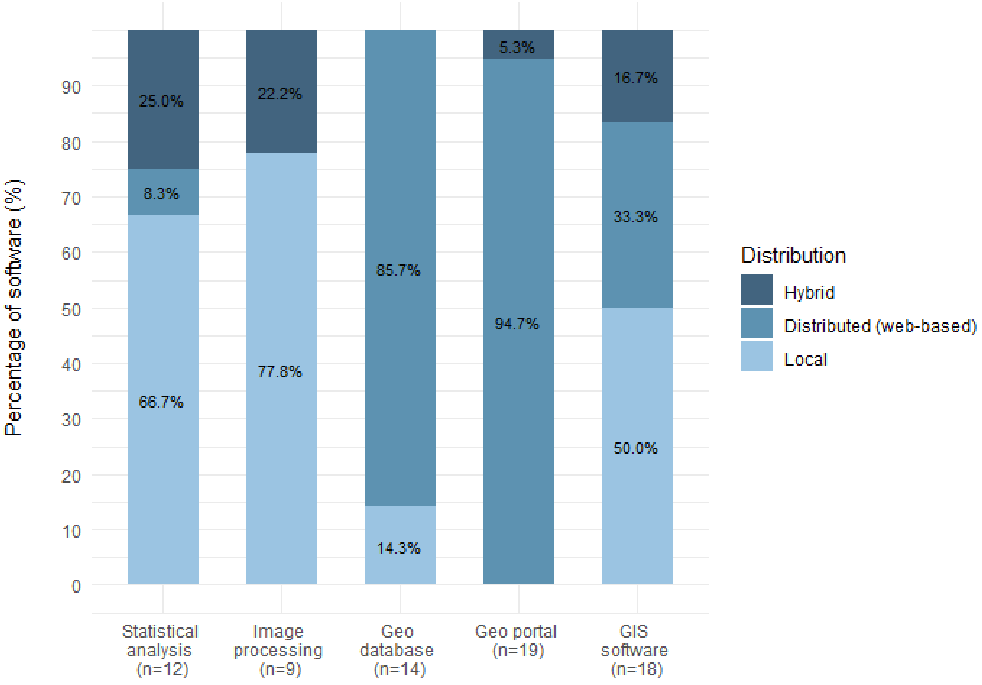


Figure 9: Type of distribution of each software type mentioned in the reviews

Web-based mapping applications are on the rise. This can be attributed to the emergence of many application programming interfaces (APIs), e.g. Google Maps API, ArcGISWeb APIs, and Bing Maps API, which allow users to build interactive web-based mapping tools (Kirby et al., 2017). Most of these mapping applications have been restricted to data visualisation, because of concerns for data protection and limited computational capabilities, although recent initiatives are in place to incorporate spatial and spatio-temporal analytical capabilities over the Web (Kirby et al., 2017).

Web GIS is a type of platform linking existing GIS systems together in a new architectural pattern, which can be used for data collection as well as for the dissemination of results (Delgado et al., 2019; Franch-Pardo et al., 2021). Agricultural Collaborative Research Outcome System (AgCROS), although still in early stages of implementation, can be considered an example of such a platform (Delgado et al., 2019). A move to WebGIS can help reduce data siloes by coalescing data sources, because it provides relatively inexpensive storage capacities (Delgado et al., 2019). However, storing data remotely may lead to legal issues regarding confidentiality (Kirby et al., 2017).

3.3.2.7. Software licensing and access

Licensing is a legal instrument that provides legally binding guidelines governing the use or distribution of software (Lutkevich, 2021). For the purpose of this report, we distinguish proprietary software from open source software. Proprietary license, also known as closed source provide operational code but restrict users from altering the software (Lutkevich, 2021). Proprietary software is copyright protected both in source code and object code forms (Wikipedia, 2022). It includes non-commercial license, proprietary license and trade secret (Wikipedia, 2022). Open source software means that the program's source code is freely available to the public (TechTerms, 2008). This type of licensing affords the user authority to modify the software functions and freedom to inspect the software code (Snyk, 2022).

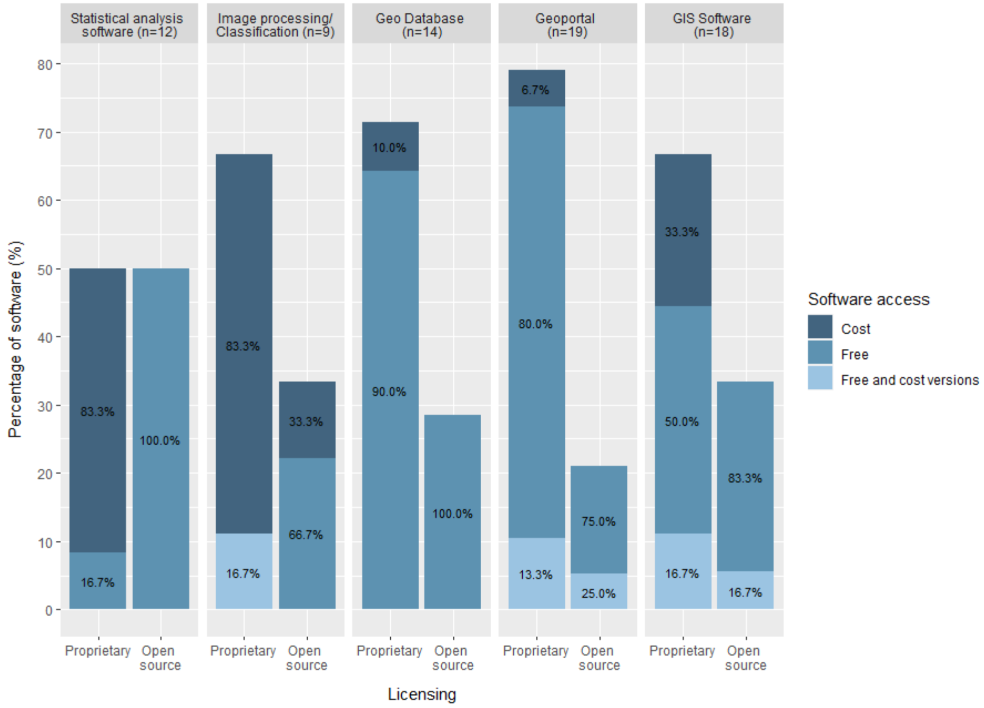


Figure 10: Software licensing and access characteristics for each software type mentioned in reviews

Proprietary software is more common, although open source solutions are becoming increasingly popular (Seidel et al., 2018; Figure 10). This can be partly attributed to the efforts of The Open Source Geospatial Foundation (OsGEO) which maintains and develops open source software packages for GIS analysis such as QGIS and GRASS (Dietrich et al., 2018; Seidel et al., 2018). While most statistical, image processing and GIS software that were identified in the reviews are proprietary and require paid licences, the majority of geoportals and geodatabases (80 and 90% respectively), although also mainly proprietary, are accessible for free (Figure 10).

3.4. Methods

A researcher can arrive at results using a multitude of methods that also vary in sophistication (Sahar et al., 2019). This section presents the scoping review findings regarding the methods used for data collection, data processing and data analysis in spatial analysis.

3.4.1. Data Collection

The development of technology (e.g. mobile devices, wearable sensors, remote sensing, satellite imagery) and the advances in GIS (e.g. web geospatial infrastructure) have allowed the collection of data at finer time and spatial resolution, while also enabling novel techniques for data collection. The findings from the scoping review show the three most mentioned methods for data collection: Ground truth data (38.5%), remote sensing (23.1%), and crowdsourcing (19.2%).

3.4.1.1. Ground truth data

Ground truth data has been identified as the most common method for data collection. It refers to the data that is collected by observation and measurement (empirical evidence) through census, surveys, and disease notification data. This data is still used in combination with high-resolution satellite data due to the latter's cost and cloud cover issues in some regions (Jia et al., 2019a). It is needed for validation of earth observation data, especially natural environmental factors (Jia et al., 2019a).

3.4.1.2. Remote sensing

Remote sensing (RS) is a method for data collection at distance able to affordably collect data at a large scale, over a short time period (Dietrich et al., 2018), typically from satellites or aircraft. It allows to conduct massive measurements (Jia et al., 2019a). This method is used in the fields of epidemiology, ecology, agriculture, oceanography, geology, and archaeology (Jia et al., 2019a). In epidemiology, "it is most used to identify determinants of infectious diseases and to develop models to predict their evolution" (Dietrich et al., 2018). In environmental health studies, it is used "to monitor air pollutants (e.g. PM2.5, PM10, O3, NO2, pollens, asbestos, volcanic ash, wildfire smoke)" (Dietrich et al., 2018). In health-related behaviour studies it can also "provide better explanations for human behaviours and decision making" (Jia et al., 2019a; Wang, 2020). However, the collection of large datasets at fine spatial and time resolution requires the development of more complex models to efficiently make use of such rich data (Bergquist and Manda, 2019).

3.4.1.3. Crowdsourcing and Citizen science

Crowdsourcing (also called citizen science), is a method that allows rapid collection of information from citizens through the internet, mobile devices or wearable sensors. It is important for real-time mapping and monitoring (Franch-Pardo et al., 2021). Crowdsourcing has been mentioned by several reviews (Jia et al., 2017; Hollings et al., 2018; Kanankege et al., 2020; Franch-Pardo et al., 2021) as a method for collection of GIS data.

3.4.2. Data processing

Regarding data processing, methods have been classified into data transformations, basic geoprocessing, smoothing techniques and image processing.

3.4.2.1. Data Transformations: geocoding

Geocoding has been identified as the most used method for data transformation in GIS analysis (30.8%), followed by data linkage (11.5%), geo-referencing (7.7%), and Fourier analysis and statistical matching (3.8%).

Geocoding means transforming location information into geographic coordinates (point data) using a dataset as reference, such as street address or postal code (Kirby et al., 2017). A geocoding service allows to automatically convert descriptive information into a point data map under a geographic coordinate system (ESRI, 2021). It allows to link different data through the spatial component and provides the basis for data visualization and analysis (Sahar et al., 2019). There are initiatives to provide guidelines for geocoding like the Open Geospatial Consortium's Geocoding API SWG, whose goal is to develop a "candidate standard for a Geocoding API" (OGC, 2022a). However, according to the reviews, there are currently no standards for geocoding systems (Sahar et al., 2019). There are differences regarding the choice of the geocoding software, address model, the choice of the reference data, and the age of the data (Wilkins et al., 2017; Sahar et al., 2019). To facilitate critical appraisal of geocoding methods, Wilkins et al. (2017) recommend authors to report the address model, percentage match rate and environmental context of the study.

Issues encountered in geocoding can be:

- positional errors due to low match score or old historical data;
- biased regression coefficients;
- inflated standard errors;
- reduced statistical power (Kirby et al., 2017; Samuels-Kalow and Camargo, 2019).

To prevent privacy issues related to spatial data, geocoded data can be presented in an aggregate manner. Similarly, to protect the privacy of individuals, researchers can geomask the geocoded data (Kirby et al., 2017).

3.4.2.2. Basic geoprocessing

Geoprocessing is any GIS data operation needed for subsequent GIS analysis. It allows to manage, compare, manipulate, and analyse input geographic information resulting in an output set of data. The scoping review findings show that buffer and network analysis are the most common geoprocessing tools (Figure 11). The creation of buffer polygons around an input feature to a specified distance allows to identify an influence zone around the input feature. Its use has been found in reviews from the fields of environmental epidemiology, human epidemiology, public health and food accessibility. Network analysis is used to find distance and spatial links between two or more input features. It is used for example to find the shortest distance between two points. In this scoping review, network analysis has been mentioned in reviews from the fields of environmental and human epidemiology, ecology, and food accessibility.

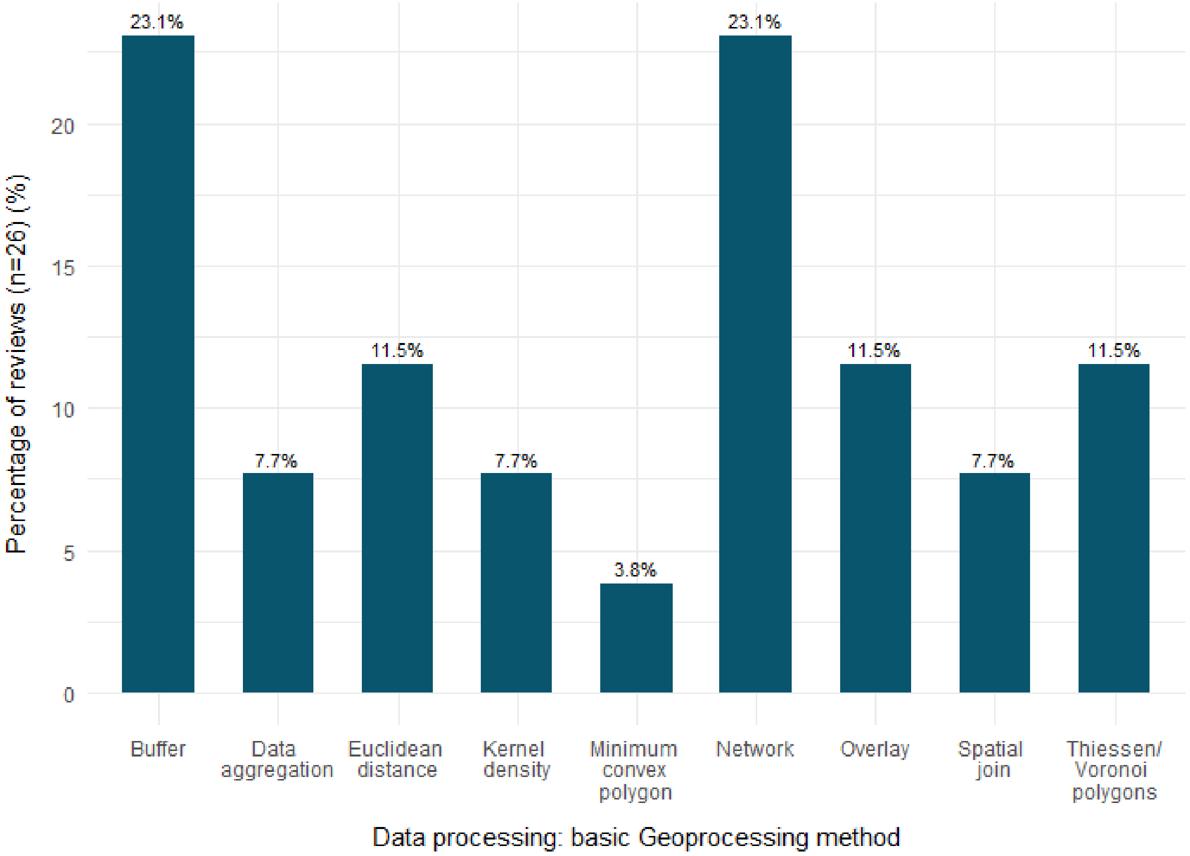


Figure 11: Percentage of reviews mentioning each basic geoprocessing method

3.4.2.3. Smoothing techniques

Smoothing is another method utilised to prepare the data for spatial analysis. Spatial smoothing techniques are applied to fill gaps in the data and to reduce random variation (Sahar et al., 2019). Common smoothing methods are kernel density estimation (KDE), kriging, and cokriging. The advantage of smoothing is that it reduces random variation which helps to more clearly demonstrate and evaluate spatial patterns (Sahar et al., 2019). However, using spatial smoothing to control the variability of small numbers can result in masking real patterns (Sahar et al., 2019).

3.4.2.4. Image processing

In GIS, image processing refers to the use of algorithms to manipulate data and process images (The GIS encyclopaedia, online). An example of image processing method can be image segmentation, such as thresholding, which can be used to distinguish objects of interest from background (Hollings et al., 2018). The scoping review findings show that supervised classification (based on supervised machine learning) appears as the most often used technique for image processing. Indeed, recent use of supervised machine learning shows drastically improved classification due to higher resolution data and significant computational capacity in today's GIS environment (Delgado et al., 2019). Unsupervised classification, whereby statistical algorithms are used to group pixels based on spectral information, has been also mentioned (Hollings et al., 2018). Unsupervised classification has been successfully used to count animals from RS imagery. However despite having high detection probability, it also led to more false positives as compared to manual counting (Hollings et al., 2018).

3.4.3. Data analysis

GIS provides the medium and the tools to conduct spatial analyses. There are two aims of analyses in public health and food safety. First aim is to identify and describe observed associations. Second is to appreciate and test the significance of the processes resulting in the observed patterns (Bergquist and Manda, 2019). Hence, "statistical methodologies are applied to describe, understand, characterize and inform about potential interventions" (Bergquist and Manda, 2019).

3.4.3.1. Spatial statistics

Spatial statistics derive from traditional statistical models and go further to integrate the spatial and temporal dimensions (Bergquist and Manda, 2019). The analysis usually involves a descriptive analysis (spatial occurrences on a map), an exploratory analysis (to search for patterns and possible explanations), a statistical analysis and hypothesis testing of expected event under assumed statistical model, and modelling and scenario analysis (involves regression models; Bergquist and Manda, 2019). Odhiambo et al., (2020) highlight the prominence and flexibility of geostatistical methods in modelling spatio-temporal disease patterns and highlight the capacity to support interpolation.

Modern mapping techniques, embracing novel statistical techniques at high spatial and temporal resolution, are increasingly being used to inform public health policy. Most recently, this has been aided by the availability of curated spatial databases, geographical information systems, enhanced computational capabilities and the advancement in spatial statistics (Cuadros et al., 2021).

This scoping review identified four most common practices of spatial analysis:

choropleth and heat mapping used for data exploration and visualisation (Bergquist and Manda, 2019, Sahar et al., 2019, Odhiambo et al., 2020, Fatima et al., 2021, Cuadros et al., 2021);

cluster analysis (Jia et al., 2017; Kirby et al., 2017; Chaix, 2018; Seidel et al., 2018; Bergquist and Manda, 2019; Broomhead et al., 2019; Card et al., 2019; Dhewantara et al., 2019; Jia et al., 2019b; Sahar et al., 2019; Kanankege et al., 2020; Odhiambo et al., 2020; Cuadros et al., 2021; Fatima et al., 2021; Franch-Pardo et al., 2021)

correlation and regression modelling (Jia et al., 2017; Kirby et al., 2017; Hollings et al., 2018; Seidel et al., 2018; Bergquist and Manda, 2019; Broomhead et al., 2019; Card et al., 2019; Dhewantara et al., 2019; Jia et al., 2019b; Malone et al., 2019; Sahar et al., 2019; Kanankege et al., 2020; Odhiambo et al., 2020; Wang, 2020; Agyeman et al., 2021; Cuadros et al., 2021; Fatima et al., 2021; Franch-Pardo et al., 2021)

Bayesian spatial statistics (Kirby et al., 2017; Seidel et al., 2018; Bergquist and Manda, 2019; Dhewantara et al., 2019; Malone et al., 2019; Sahar et al., 2019; Cuadros et al., 2021; Kanankege et al., 2020; Odhiambo et al., 2020).

Choropleth and heat mapping

Choropleth and heat maps are useful for visual exploration of data. Both are common statistical maps used for measuring variability of events across a geographic area (Bergquist and Manda, 2019). Choropleth maps are thematic maps, that use the intensity of colour to correspond to an aggregate of an event or characteristic (Bergquist and Manda, 2019). Issues to be considered when working with choropleth maps are the modifiable areal unit problem (MAUP), and the ecological fallacy, both resulting in misleading conclusions stemming from a wrong interpretation of the geographic unit (Bergquist and Manda, 2019). For a detailed description of these problems see section 3.4.5.

Heat maps are slightly different than choropleth maps. While the latter show changes across an area with geographic boundaries, heat maps represent density across continuous data (e.g. density or elevation; Bergquist and Manda, 2019). They are generally used when tracking natural phenomena or for disease mapping (Fatima et al., 2021).

Cluster analysis

Cluster analysis is commonly used to detect spatial autocorrelation among events; it tests whether the spatial distributions of similar events is more clustered than a random spatial distribution. In public health and epidemiology, cluster analysis is fundamental to identify hotspot areas which uncover the location of high-risk populations and identify the triggering factors of an epidemic (Cuadros et al., 2021). Cluster and hot spot analysis was particularly important for COVID-19 surveillance by identifying the size, location, and risks of the disease (Fatima et al., 2021). To identify and predict COVID-19 spread, space–time clustering techniques, concretely Poisson space–time scan statistic, were also used by researchers (Fatima et al., 2021). Similarly, Odhiambo et al., (2020) found that cluster detection methods were used in about 30% in malaria studies to identify significant geographical variation of the disease. Well known techniques of cluster analysis are the local and global Moran's I, Geary's C, Getis and Ord Gi, and scan statistics (e.g. Kulldorf spatial scan statistic ; Bergquist and Manda, 2019).

Correlation and regression modelling

GIS allows to link health outcomes with the geographic conditions (Bergquist and Manda, 2019). Standard correlation and regression models can be used to reveal relationships between environmental information and parameters (e.g. poverty; Bergquist and Manda, 2019). Issues appear in the presence of expected spatial correlations, where "estimated coefficients and coefficients of determination would appear larger, and standard errors smaller than they really are; this could lead to potentially incorrect conclusions regarding the statistical significance of a relationship" (Bergquist and Manda, 2019). Commonly used methods are spatial regression models, spatial autoregressive lag model error models, and geographically weighted regressions. Bivariate spatial correlation and multivariable spatial regression models where used for exposure mapping in COVID-19 studies, (Fatima et al., 2021).

Bayesian spatial statistics

Finally, Bayesian spatial statistics is also commonly used for studying events in public health and food safety. In Bayesian statistics, "probability is expressed as a degree of belief in the occurrence of an event, rather than a fixed value based on frequency" (Bergquist and Manda, 2019). There has been a growth in epidemiological applications of Bayesian spatial modelling thanks to the advances in computational power (Bergquist and Manda, 2019). Bergquist and Manda (2019) have identified the most common Bayesian models as:

Bayesian spatial modelling;

Bayesian estimation and inference of parameter estimates, model specifications and comparisons as well as predictions.

Odhiambo et al. (2020) identified that Bayesian spatio-temporal methods were used to:

- Predict risk at unsampled locations;
- Improve predictions in areas with high variation of observations;
- Model the spatial and temporal correlations of disease occurrence, morbidity, etc.;
- Link environmental data and outcome of disease.

In health and epidemiology, Bayesian spatial modelling has gained preference over conventional classical spatial modelling (Bergquist and Manda, 2019). Applications of Bayesian spatial methods have been mostly based either on the lattice geographical units model of Besag, York and Mollie (BYM), the convolution model, or the distance-based geostatistics model advanced by Diggle and Ribeiro (Bergquist and Manda, 2019).

3.4.4. Benefits of spatial analysis

Spatial statistics enable creating simulations and conducting scenario analysis. Thus, GIS in combination with spatial statistics helps researchers to transition from description to prediction. In epidemiology, disease and outbreak prediction is crucial to design control measures against communicable and non-communicable diseases such as ischemic heart disease, pancreatic cancer, and leukemia (Dhewantara et al., 2019; Malone et al., 2019; Kanankege et al., 2020; Odhiambo et al., 2020; Cuadros et al., 2021; Fatima et al., 2021).

Applications of novel geo-statistics allow, among other, to predict risk for invasive plants and animals, conduct climate suitability analysis and assess how can established species cope with climate change (Bergquist and Manda, 2019). Also, in the context of agriculture and soil management, spatial statistics have been applied to predict potentially toxic elements in soil and sediments.

Other advantages of spatial statistical methods are:

- Identifying geographical hotspots and disease burden determinants that may facilitate design and implementation of control measures (Cuadros et al., 2021);
- Identifying high-risk populations;
- Revealing the factors that facilitate persistence and spread of epidemics;
- Detecting disease evolution in space over time;
- Helping to understand mapped information by identifying patterns and drivers of disease distribution;
- Elucidating the spatial distribution of population health to guide intervention capacity (Fatima et al., 2021)

3.4.5. Challenges of spatial analysis

Several challenges have been identified imbedded in spatial analysis:

- Modifiable aerial unit problem (MAUP)
- Ecological Fallacy
- Cartographic confounding
- Small Population Problem
- Spatial misclassification or "uncertain geographic context problem (UGCoP)"
- Complexity of methods
- Methodological variability and lack of guidelines

3.4.5.1. Modifiable aerial unit problem (MAUP)

This problem happens when using different unit of analysis leads to different geographic analysis' conclusions. To reduce the MAUP, some empirical approaches attempt to draw geographic units in non-arbitrary administrative boundaries (Kirby et al., 2017). Some researchers have even gone further to propose the creation of new synthetic areas for GIS analysis rather than existing administrative units (Samuels-Kalow and Camargo, 2019). For example, following time geography, recent studies use GIS to track individual's exposure to environmental risk factors along a time continuous by tracing mobility using portable sensors (Wang, 2020).

3.4.5.2. Ecological Fallacy

The problem arises when the results obtained from aggregated data are assumed to apply to individual cases (Bergquist and Manda, 2019). The use of smaller geographic units of analysis in maps can decrease the risk of ecological fallacy and MAUP, but maps can also become less readable (Bergquist and Manda, 2019).

3.4.5.3. Cartographic confounding

Phenomenon that happens by the risk of low match rates leading to unrepresentative measures of a given metric, or match rates that vary geographically in a non-random manner possibly leading to confounded results (Wilkins et al., 2017).

3.4.5.4. Small Population Problem

Public health data analysis often suffers from the small population problem. According to Wang (2020), it is caused by "less reliable rate estimates, sensitivity to missing data and other data errors, and data suppression in sparsely populated areas". Accordingly, the problem is more evident in rare disease cases or by population subgroups.

3.4.5.5. Spatial misclassification or "uncertain geographic context problem (UGCoP)"

Spatial misclassification refers to the use of areal units that do not accurately reflect "the true exposure" area (Wang, 2020). The uncertain geographic context problem (UGCoP) refers to the methodological question of identifying an areal unit of analysis that accurately captures the relevant geographic context instead of relying on pre-defined administrative units (Wang, 2020). Geographic momentary assessment helps to deal with the spatial misclassification problem as the areal unit is not limited to a static predetermined administrative boundary but is opened to a spatio-temporal continuum that better defines contextual environments (Chaix, 2018; Korycinski et al., 2018).

Several geographic strategies have been attempted to mitigate the problem (constructing representative geographic areas) with spatial smoothing methods (Wang, 2020):

- floating catchment area method
- kernel density estimation
- locally weighted average
- adaptive spatial filtering

3.4.5.6. Complexity of methods needed to process GIS data

As data collection techniques and computational power advance, spatial analysis also becomes more elaborated. Hence, the complexity of spatio-temporal models has increased, making inferential and predictive processes also more difficult to undertake (Odhiambo et al., 2020). Furthermore, the difficult computational algorithms and complex spatio-temporal data structures, together with limited technical expertise, may limit their applicability

(Odhiambo et al., 2020). Therefore, Jia et al. (2017) recommend to "develop friendly data processing and analysis methods to enable more researchers to process, combine and analyse GIS and other types of data".

3.4.5.7. Methodological variability and lack of guidelines

Reviews have identified variability among the methods and lack of guidelines for statistical methodology, which hinders the interpretation and reproducibility of scientific findings (Wilkins et al., 2017; Odhiambo et al., 2020). Given the heterogeneity of methods, accurate and transparent reporting of methodology is essential (Wilkins et al., 2017). However, Wilkins et al. (2017) also recognize that there is little guidance to support authors in reporting GIS methods. Given this challenge, Bergquist and Manda (2019) agree that there is the urgency to meet "a consensus on the complexity of spatial models used for estimating and predicting spatial surfaces in order to provide robust and reliable spatial mapping".

3.4.6. Methods by metric

The following section presents the findings on methods by the metrics of movement, spread, accessibility and exposure. These metrics provides frameworks

3.4.6.1. Movement (animal movement patterns, time geography, momentary assessment)

This report considers movement in association with exposure, time geography and prediction. Techniques and applications of this metric are extensive, but here we provide some examples as found in the reviews.

Studying the movement of animals is relevant for epidemiology as it reveals possible disease spread. In animal ecology, different analyses are applied to study the movement of animals (path-based metrics, wavelet analysis, recursion analysis, change-point-analysis, contact network, etc.), and varying behaviour patterns across temporal scales. Finally, cluster analysis is used for grouping trajectory points (or segments) that represent the same behavioural state or syndrome (Seidel et al., 2018).

Momentary analysis also looks at individuals' movement across a time and space continuum to assess exposure or behaviour (Wilkins et al., 2017; Chaix, 2018). Momentary analysis benefits from a continuous time-space data resolution which is made possible thanks to the advancement of data collection technology like sensors (Wilkins et al., 2017). Methods for momentary analysis include spatial interpolation and algorithms for mode recognition, and machine learning based on GPS data (Chaix, 2018).

3.4.6.2. Spread

The metric of spread is used in public health for assessing disease patterns mapping and prediction. GIS enables to combine the geographic data of a disease distribution, like COVID-19, with additional influencing factors, like demographic or environmental factors (Fatima et al., 2021).

Probability of disease spread can be measured with spatial regression models (spatial autoregressive, conditional autoregressive, geographically weighted regression) if the data has a spatial component (Kanankege et al., 2020). To measure probability of disease spread with point pattern data, point process models are used (e.g. conditional logistic model, separable and non-separable models for spatio-temporal data; Kanankege et al., 2020). Spread is also measured by Plume models and the Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPLIT). However, both models require costly high computational power due to their complexity (Kanankege et al., 2020).

3.4.6.3. Accessibility

Accessibility is another recurrent metric in the domain of food safety and public health. It is used to assess the "ease of reaching a venue" (Jia et al., 2017). Applications include measuring accessibility to health services, food outlets, and retail food environments (Jia et al., 2017; Kirby et al., 2017; (Wilkins et al., 2017); Broomhead et al., 2019; Jia et al., 2019b; (Kost, 2019)); Samuels-Kalow and Camargo, 2019; Wang, 2020; Cuadros et al., 2021). Further, accessibility assessment informs decision taking and policy making regarding resource allocation, based on potential crowdedness of facilities (Jia et al., 2017; Bergquist and Manda, 2019; Broomhead et al., 2019; Card et al., 2019; Jia et al., 2019b, Wang, 2020; Cuadros et al., 2021).

However, the metric of accessibility faces two challenges discussed above; the modifiable areal unit problem ((Wilkins et al., 2017) and the small population problem. Also, there are no standards for defining the area of study, for example area of food outlet cover (Wilkins et al., 2017; Wang, 2020), which leads to the 'uncertain geographic context problem (UGCoP; Wang, 2020). When the contextual unit is ill defined, findings can be affected by ecological fallacy.

Common techniques found in accessibility measurements are overlay analysis (Jia et al., 2017; Malone et al., 2019; Kanankege et al., 2020), buffer analysis (Jia et al., 2017; (Wilkins et al., 2017); Broomhead et al., 2019; Jia et al., 2019b; Malone et al., 2019; Kanankege et al., 2020; Wang, 2020; Cuadros et al., 2021), network analysis (Jia et al., 2017; Seidel et al., 2018; Card et al., 2019; Jia et al., 2019b; Malone et al., 2019), Euclidean distance (Card et al., 2019), and the 2-step floating catchment area (2SFCA) by Luo and Wang (Wang, 2020).

3.4.6.4. Exposure

Exposure is measured in environmental health, and neighbourhood research to assess environment's effect on individuals' health (Chaix, 2018). It does so by looking at individual's exposure to environmental conditions and events across time and space. The exposure mapping is also used in spatial epidemiology to measure individuals' exposure to certain risk factors that may trigger a disease (Korycinski et al., 2018; Bergquist and Manda, 2019; Kanankege et al., 2020; Odhiambo et al., 2020; Cuadros et al., 2021; Fatima et al., 2021). Exposure metrics are also used in health-related behaviour studies (Wang, 2020), life-segment momentary analysis (Chaix, 2018), obesity research where exposure to obesogenic features is measured (Jia et al., 2019b), cancer control and population sciences (Korycinski et al., 2018), malaria risk (Odhiambo et al., 2020), COVID disease pattern detection (Fatima et al., 2021), disease and exposure mapping and modelling (Fatima et al., 2021), and non-communicable disease (diabetes) monitoring (Cuadros et al., 2021). Lastly, improved exposure measures on daily mobility has benefited from smartphone-based GPS tracking, modern GIS analytics, and environmental sensing (Wang, 2020).

There are following challenges when assessing exposure:

- Defining areal units of analysis across a time and space continuum
- Modifiable areal unit problem
- The small population problem
- Failure to consider cumulative exposures over time (e.g., residential history; Chaix, 2018).

Finally, exposure variables are becoming more algorithm-based (Chaix, 2018). However, researchers must assess the efficiency of these to predict (a) perceptions of the corresponding exposures or (b) behavioural/health outcomes in comparison to more rudimentary variables (Chaix, 2018).

Key messages

- GIS has increasing presence in public health and food safety given the rising awareness of geographical and temporal dimensions of health
- GIS allows a holistic look at events thanks to its ability to join data across disciplines and geographic boundaries.
- Spatial analyses conducted with GIS facilitate data exploration and interpretation.
- Attention in public health and food safety is turning from control to prevention. GIS supports this transition as it provides the medium to conduct spatial statistical analyses that allow to move beyond description to simulation.
- Public data sources were found more common.
- New methods and technology advances allow to collect data at increasingly finer temporal and spatial resolution.
- Advances in data collection allow the continuous collection of data, which enables expanding the study of health at a spatio-temporal continuum.
- The immense amounts of data collected at increasingly finer resolutions pose challenges regarding the use, sharing and analysis of GIS data.
- There is a trade-off between data resolution and cost.
- Processing and analysing large amounts of data implies more complex models, difficult to undertake and interpret.
- Finer scale implies more policies related to geoprivacy, which limits data shareability and publication, especially when human beings are involved.
- For conducting spatial analysis, there is a general preference for using specialised GIS software (such as ArcGIS, QGIS), and software environment and programming languages for statistical computing that support spatial analysis (R), followed by geoportals.
- There is a high presence of web-based applications (85% of geodatabases and geoportals had a web-based distribution) and the use of cloud computing, which allows the rise of a Web GIS that allows to integrate data thanks to the internet, creating a network of data and overcoming data siloes. However, the storing and sharing of data online poses legal issues regarding confidentiality.
- Proprietary software is more common, although open source solutions are becoming increasingly popular.
- Despite initiatives as INSPIRE Directive and the Destination Earth (DestinE) initiative, there are still not enough standardized practices nor models to process and analyse data, and spatial data is still limited by data fragmentation and siloes, because of competing interests and variety of formats.

4. Visual exploration

This report includes a visual exploration of the uses of GIS by public agencies and Supra National Bodies[13] from the domain of public health and food safety. The visual exploration consists of exploring the selected agencies' websites with the aim of finding any evidence of the use of GIS in these domains. This visual exploration was conducted in parallel with the scoping review, to complement the findings from the literature, with concrete examples on how and why the agencies use GIS.

The visual exploration looked at the websites of the following agencies from the public health and food safety domains from Europe and USA[14]:

Europe:

- European Centre for Disease Control (ECDC)
- Environmental European Agency (EEA)
- European Statistical Office (EUROSTAT)
- European Commission (EC)
- Joint Research Centre (JRC)

USA:

- U.S. Food and Drug Administration (FDA)
- U.S. Centers for Disease Control and Prevention (CDC)
- U.S. National Cancer Institute (NCI)
- U.S. Department of Agriculture (USDA)

Besides, this section includes information retrieved from two summits: GIS EU virtual summit 2020 and GIS EU virtual summit 2021.

The findings capture a general view into the current GIS scenario in the fields of human, plant and animal health, agriculture, and environment. It is important to note that given the variety of missions and activities of the explored agencies, this report presents rather a general picture of GIS activities. These findings however point out to the general trends and serve as an initial basis for future in-depth research of the GIS applications in the field of food safety and public health.

4.1. Fields and applications of GIS

The visual exploration identified that GIS is applied in many fields (see Table 23, Appendix I for a complete list of fields). However, we would like to highlight the use of GIS in the following fields:

- Human epidemiology and Public health
- Agriculture
- Plant and Animal health
- Environment

The applications of GIS were identified for each of these fields. Regardless of field, public agencies use GIS to:

- Conduct analysis
- Facilitate communication
- Support and inform policy makers

4.1.1. Conducting Analysis with GIS

---

[13] In Chapter 4, supra national bodies and public agencies will be collectively referred to as agencies.

[14] See Table 24, Appendix II for a list of the agencies' websites with the corresponding link

GIS is used to conduct analysis, with a focus on studying health in relation to space, time and the environment. The following section exemplifies the analysis that can be conducted with GIS' support by fields of application.

4.1.1.1. Human epidemiology and Public health

GIS is used by the ECDC (Geoportal, Geocatalogue and map maker tool EMMa), CDC (all divisions included in Table 24, Appendix II) and NCI (GIS portal for cancer research) to:

- Monitor incidence, outbreak and distribution of diseases
- Visualise epidemiological data
- Improve disease detection, prevention and response capacity
- Assess spatio-temporal relationships of disease and climate factors and environment
- Support communication and sharing of scientific findings

4.1.1.2. Agriculture, Plant and Animal health

The USDA, and its divisions for Animal and Plant Health Inspection Service (APHIS) and Global Agricultural & Disaster Assessment System (GADAS) use GIS to:

- Assess the agricultural impact of severe droughts and other natural disasters
- Regulate introduction or dissemination of plant pests
- Perform site suitability analysis, cluster/hotspot identification
- Collect field data of plants / pests
- Improve pest detection, prevention and response capacity
- Perform pest outbreak surveillance
- Conduct risk assessment of disease introduction to agricultural animals
- Respond to animal disease outbreaks
- Support emergency response

4.1.1.3. Environment

GIS is used by agencies from the environmental sector (EEA Disco Map, SDI – geospatial data catalogue, SIIF - UWWTD) and by agencies that analyse a given field (e.g. health) in relation to the environment, such as the ECDC, U.S. CDC (all divisions listed in Table 24, Appendix II) and USDA (APHIS) to:

- Integrate data across different disciplines and geographic borders to holistically address a global environmental crisis
- Study the relationship of human, animal and plant health and the environment
- Manage the urban water system

4.1.2. Facilitate communication

The support of GIS on conducting analysis has been complemented by its capacity to communicate. The visualisation of spatial data in maps and dynamic visualisations such as dashboards and story maps help not only identifying previously covered trends, but also efficiently communicating those findings to policy makers. The following examples show how agencies have recognized the usefulness of GIS to communicate:

- ECDC's EMMa: "maps make it much easier to identify and illustrate trends" (ECDC, 2022).
- CDC's Division for Heart Disease and Stroke Prevention – NCCDPHP: "interactive graphics and maps […] provide visual support for deciding where to focus cancer control efforts". This "quick and easy access to descriptive cancer statistics" allows to "prioritize investments in cancer control" (CDC, 2020b).
- CDC's Geospatial Research, Analysis, and Services Program (GRASP): "maps and geospatial resources […] are essential to CDC/ATSDR's efforts to communicate complex results to communities, government partners, academia, and advocacy groups" (CDC, 2020a).

- USDA APHIS: story maps are used to “share information with the public, stakeholders and customers […] by providing interactive and engaging content” (APHIS).
- In short, GIS provides the capacity to visualise and contextualize data, which allows to illuminate phenomena previously obscured and communicate the findings in appealing and evident ways.

4.1.3. Support policy making

Given the previously mentioned benefits, GIS helps guiding policy making. Specially under the increasing attempts to tackle the current climate and sanitarian crisis holistically, as seen with the ONE Health Concept, GIS plays a crucial role integrating data and organisations across disciplines and borders, which helps informing better policies (EsriEUSummit, 2021). Different agencies using GIS reiterate the importance of spatial data and the geospatial infrastructure for the following reasons:

It allows better understanding of the context across different topics (environment, economy, etc.), which leads to better decisions and policy making.

- It enables to integrate data at the regional level for holistic policy making.
- It helps to translate information to knowledge thanks to visualisation.
- It informs disease and outbreak prevention programs and policies.
- It identifies resources and communicates needs to decisions makers.

4.2. Policy background for the use of GIS

The implementation of GIS has evolved throughout the years with policies advocating for:

Shareability and reuse of data:

- INSPIRE Directive 2007/2/CE: on the harmonization and shareability of spatial data across Europe

Availability and accessibility to open data:

- 2003/4/EC - on public access to environmental information15
- Directive 2003/98/EC to promote accessibility and reuse of Public Sector Information (PSI)16
- EU open data strategy on a single European data space, data sovereignty, availability and reusability
- Directive (EU) 2019/1024 on open data and the re-use of public sector information17

Standardization, harmonization of datasets:

- COM(2022) 68 European Data Act to lay down a harmonised framework for an internal data market18

Centralized data repositories:

- COM (2008) 46 on the simplification of the collection, exchange and use of data based on access, sharing and interoperability19
- COM(2022) 31 on the EU strategy on data standardization of a single digital EU market20

---

[15] Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC

[16] Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information

[17] Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information

[18] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on harmonised rules on fair access to and use of data (Data Act), p.17

[19] Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions - Towards a Shared Environmental Information System (SEIS)

[20] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS An EU Strategy on Standardisation Setting global standards in support of a resilient, green and digital EU single market, 2022

Cost-efficient collection and integration of data:

- The Geospatial Data Act of 2018 (GDA) to ensure the availability, interoperability and affordability of integrated data21

These policies stand around the European strategy for data, which sets requirements for open and secure data sharing, and shared data spaces for business and public sector collaborations. At the same time, this wide use of data is balanced with privacy, security and high ethical standards (Saligoe-Simmel in EU Esri Summit, 2021).

4.3. Geospatial infrastructure

The visual exploration has revealed two recurrent geospatial infrastructure types:

- Geoportals
- Geodatabases

4.3.1. Use and benefits of geospatial infrastructure

Geoportals and geodatabases are a crucial part of a web-geospatial infrastructure. Agencies are increasingly aspiring to implement a geoportal or geodatabase to:

- facilitate data access and use (INSPIRE Geoportal, ECDC Geoportal, Geocatalogue and EMMa, EC Data Europa, Eurostat GISCO, EEA SIIF-UWWTD, EEA SDI - geoportal data catalogue)
- support data interoperability (EC Data Europa, CDC - all divisions listed in Table 24 , Appendix II)
- allow interconnection of data across different scales, sectors, geographic boundaries and disciplines
- provide centralized data repository (EC Data Europa, INSPIRE Geoportal)
- allow the visualisation of spatial data in a simple and straightforward way, overcoming former technical barriers for using GIS (ECDC Geoportal, Geocatalogue and EMMa, EEA Disco Map, CDC Atlas Plus, USDA GADAS)

Self-service web geospatial infrastructure (such as geoportals) provide services to conduct exploration and general geospatial analysis that help to overcome technical barriers towards the use of GIS. Some go even further to provide spatial statistics services (e.g. Eurostat GISCO, CDC State Cancer Profiles, CDC U.S. Cancer Statistics Data Visualization, CDC Epi Info™).

However, it is important to consider that while the choices for accessing spatial data increase, and therefore, there are more options for retrieving data as seen in Inspire Geoportal and Data.Europa, the official portal for European Data, the process of searching and retrieving data also becomes more complex. Hence, centralized data repositories are important to ease the process of data sharing and retrieval. Well documented metadata is also important for facilitating the retrieval of data. Examples of geodatabases with well-documented metadata are: INSPIRE Geoportal, European Data Portal, ECDC Geo Catalogue, and the EEA Geospatial Geo-catalogue, because they provide a thorough description of the dataset, and indicate the temporal and spatial extent, and include technical information such as the spatial and temporal resolution, format and the coordinate reference system.

4.3.2. Type of GIS application provided

The analysis has identified that all the explored agencies use GIS to support conducting spatial analysis, with about 81.8% providing geoportals or geodatabases (Figure 12).

---

[21] USA Federal Geographic Data Committee, Geospatial Data Act of 2018

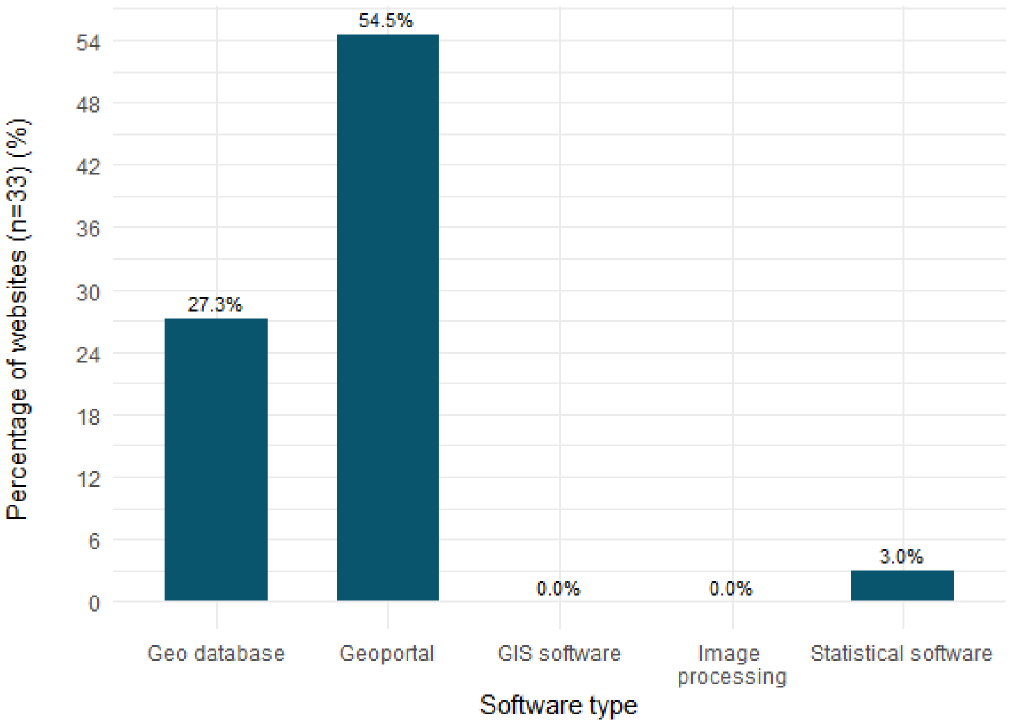


Figure 12: Percentage of websites mentioning each software type

The visual exploration elucidated that most of the agencies provide public access to geospatial applications. Almost all explored geoportals and geodatabases (more than 90%) offer open access to the provided services (e.g. visualisation or discovery services). However, access to data is often limited (Figure 13**Error! Reference source not found.**). From the open software, 44 % provide open data access, while 40% give limited access to the data due to "protection of international relations, public security or national defence" (INSPIRE Directive, 2007).

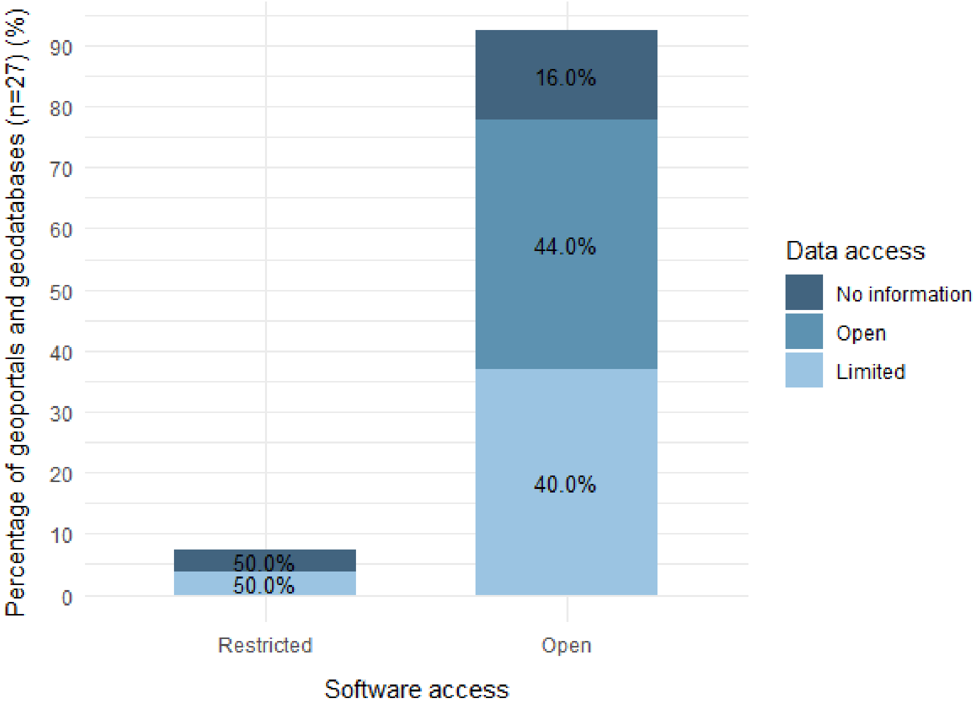


Figure 13: Software access[22] and data access characteristics for the geoportals and geodatabases mentioned in the websites

Key messages

- GIS is used by agencies in the fields of human epidemiology and public health, agriculture, plant and animal health, environment, among others.
- GIS is used by the explored agencies to 1) conduct analysis, 2) facilitate communication, and 3) support and inform policy makers.
- Geoportals and geodatabases are a common GIS infrastructure located in the web. The advantages of these are:
    - facilitating data access, use, and visualization
    - supporting data interoperability
    - allowing interconnection of data across different scales, sectors, geographic boundaries and disciplines
    - providing a centralized data repository
    - providing self-services to conduct exploration and general geospatial analysis that help to overcome technical barriers
- There is an urgent need for centralized access to data. There are more and more portals and options for data access, which on the one hand, support data shareability, while at the same time complicate searching for data.
- Well-documented metadata is crucial for facilitating data finding and reusability.

5. GIS at EFSA

To guide future implementations of GIS at EFSA, it is necessary to track the current GIS scenario at the agency. The use of GIS is increasingly present in the scientific fields of public health and food safety, and it is more and more used at EFSA for different purposes. Therefore, there is the need for a more comprehensive view of the ongoing GIS activities and projects.

Given the relevance of the topic, an internal survey was circulated to identify the activities and projects using GIS at EFSA, and the corresponding staff in charge.

The survey was circulated in two stages. The results of the first survey captured a robust snapshot of the current scenario at EFSA. However, these results could not be considered representative given the small sample size. Hence, the survey was circulated a second time with the whole EFSA scientific community. The second circulation of the survey did not affect the results of the first survey but helped confirming the number of EFSA staff (6.25%) using GIS.

The survey results helped identifying an external project conducted by EFSA in collaboration with other European bodies. This case is presented at the end of chapter 5, as part of EFSA's activities but conducted beyond EFSA's internal scope.

---

[22] Software access is here defined as the availability to access the software and the services provided by this. It can provide a free access to the software and services, or have a restricted access, dependent of a payed license.

5.1. Descriptive results: sample and number of responses

The first survey was circulated with 22 staff that were previously identified as GIS users. A total of 19 staff completed the survey, from which 14 were currently applying GIS techniques at work. The second survey was circulated within the EFSA scientific community, a total of 240 staff. From these, additional 12 answered the questionnaire, and only 2 apply GIS techniques at work (Table 6).

Table 6: Survey responses

Survey	Sent	Received	GIS users
1st circulation	22	19	14
2nd circulation	240	12	2
Total	240	31	16

Survey responses were received from 12 EFSA's units, mainly BIOHAW, PLANTS, MESE, and iDATA (Table 7). GIS is mostly used by EFSA's scientific units (88%; BIOHAW, PLANTS and MESE), and it is mostly applied in biological monitoring and events that require spatial awareness. GIS technology is also used by iDATA and MESE to manage spatial data and develop methodology to give scientific support to EFSA's scientific units.

Table 7: Survey respondents by Units and GIS users

UNIT	Responses	GIS users	GIS users in %
BIOHAW	8	6	38
PLANTS	3	3	18
MESE	7	3	18
iDATA	3	2	14
PREV	2	–	–
KNOW	2	1	6
NIF	1	1	6
CORSER	1	–	–
LA	1	–	–
FEEDCO	1	–	–
FDP	1	–	–
FIP	1	–	–
Total	31	16	100%

–: no answer

5.2. Theoretical Section

This section explores how and in which fields GIS is used at EFSA. It also enquired about the methodology used to conduct any GIS activity, and the legal implications thereof.

5.2.1. GIS use at EFSA

GIS is mostly used at EFSA for biological monitoring, especially in the fields of animal health and welfare, and plant health. In animal health, GIS is mostly used for the epidemiological assessment of African Swine Fever, while in plant health, it is mainly used for conducting plant and pests risk assessment, climate suitability analysis, and crops distribution survey.

GIS is used in biological monitoring mainly to:

- Describe the geo-distribution of an event (single cases of disease, abundance of a given vector / insect / animal / plant in an area).
- Simulate and identify potential (geographical) risk factors.
- Support decision taking and policy making.

The following tables (Table 8; Table 9; Table 10; Table 11)[23] present the main activities where GIS is applied by field.

Table 8: GIS activities in the field of animal health and welfare

Field	Description of activity
Animal Health Law (ASF, AI, other)	Elaboration of maps to describe geographical distributions of certain entities
	Use of story maps for interactive presentation of EFSA outputs
African Swine Fever (ASF; epidemiological assessment)	Depiction of distribution of ASF outbreaks and wild boar/pig population
	Evaluation of risk factors related to the occurrence of an ASF outbreak
	Demonstration of prevalence and distribution of ASF outbreaks in time at Nomenclature of Territorial Units for Statistics (NUTS) level
	Development of spatio-temporal models with spatial explicit information to identify risk factors
	Definition of potential secondary cases associated to a primary case in the spread of a disease
	Production validation dashboards with ASF data
	Development of applications that use GIS information to produce maps that are used in risk assessments in ASF
	Development of network models to replicate spread
	Publication of Scientific Opinions on ASF (annual epidemiological report and other mandates)
Avian Influenza (AI)	Risk assessment on avian influenza in and outside the EU on a quarterly basis
Description of disease profiles	Describe outbreak location, and spread of vector-borne diseases
Animal population	Collection of Pig and wild boar / poultry and wild birds data
Animal Ecology	Model the presence, distribution and abundance of wild boar and other mammals
Bee health	Visualization of bee health data to support communication and decision taking related to pollination

ASF: African Swine Fever; AI: Avian Influenza

Table 9: GIS activities in the fields of zoonoses (human and animal) and antimicrobial resistance (AMR)

Field	Description of activity
Zoonoses	Production of joint scientific reports EFSA-ECDC on the EU One Health Zoonoses
	Elaboration of joint EU maps with several outcomes on human and animal health
	Processing and management of the geographical dimension of monitoring data received from MS
Antimicrobial resistance (AMR)	Assessment and presentation of the occurrence of resistance of different combination of antimicrobials and bacteria in Europe

Table 10: GIS activities in the field of Plant health

---

[23] Tables 8 – 11 contain information as formulated by the respondents.

Field	Description of activity
Plant health	Assessment of the potential risk of introduction of pests in plant and crops production areas
	Determine area of potential establishment of plant pest and diseases
	Use of maps to identify potential points of pest entry
	Use of maps to represent pests' spread and survival
	Production of maps to simulate pests outbreak
	Visualize geographical distribution of pests and their hosts

Table 11: GIS activities in other fields

Field	Description of activity
Sustainable food production	Consider the geo-distribution and management of crops and non-target species, to assess the geo-distribution of exposure and risks in relation to other sustainability indicators.
Genetically modified organism (GMO)	Development of a spatially and temporally explicit model to quantify risks to non-target Lepidoptera associated to the ingestion of genetically modified Bt-maize pollen deposited on their host plants.

GMO: Genetically modified organism

5.2.2. Methodology

This section collected the methodology used by the different activities involving GIS at EFSA. The activities and projects conducted by different fields use a varied methodology that is presented in

Field	Sub-field	Methodology
Animal Health Law	African Swine Fever (ASF)	Visualization using story maps and choropleth maps
		Descriptive epidemiology
		Geo-processing data
		Use of the spatial tool from R4EU for ASF annual reports
		Spatio-temporal models, Network models, GLM, and others
		Visualize geographical distribution of pests and their hosts
	Description of disease profiles	Generalized Linear Models (GLM), Besag-York-Mollié (BYM)
	Animal Population	Collection of data of pigs and wild boars and poultry and wild birds
	Animal Ecology	Presence-background or presence-absence modelling approach in addition to a presence-only approach
	Bee Health	Algorithms developed by contractor
Zoonoses and AMR	Zoonoses	Spatial join between zoonoses data and country-level or sub-country level using NUTS codes
	Antimicrobial Resistance (AMR)	Mapping of occurrence of resistance
Plant Health		Buffer analysis and maps overlay for High Risk Plants
		Climate suitability modelling based on biophysical parameters
		Overlapping of data in order to determine areas of suitability and/or impact

[24]. Among the identified activities, several use GIS to:

- Conduct descriptive analysis or visual exploration using general geoprocessing methods such as buffer, overlays and spatial join (Zoonoses, ASF, Plant health, AMR).
- More complex spatial models have been designed for animal health monitoring (ASF) and plant health (plant/pests risk assessment and climate suitability).
- Spatial statistics are used to describe and model disease profiles (vector-borne diseases). General linear models (GLM) and the Besag York Mollié Model (BYM) have been used for this purpose.

---

contains information as formulated by the respondents.

Table 12: GIS methodologies used by EFSA in different fields

Field	Sub-field	Methodology
Animal Health Law	African Swine Fever (ASF)	Visualization using story maps and choropleth maps
		Descriptive epidemiology
		Geo-processing data
		Use of the spatial tool from R4EU for ASF annual reports
		Spatio-temporal models, Network models, GLM, and others
		Visualize geographical distribution of pests and their hosts
	Description of disease profiles	Generalized Linear Models (GLM), Besag-York-Mollié (BYM)
	Animal Population	Collection of data of pigs and wild boars and poultry and wild birds
	Animal Ecology	Presence-background or presence-absence modelling approach in addition to a presence-only approach
	Bee Health	Algorithms developed by contractor
Zoonoses and AMR	Zoonoses	Spatial join between zoonoses data and country-level or sub-country level using NUTS codes
	Antimicrobial Resistance (AMR)	Mapping of occurrence of resistance
Plant Health		Buffer analysis and maps overlay for High Risk Plants
		Climate suitability modelling based on biophysical parameters
		Overlapping of data in order to determine areas of suitability and/or impact

5.2.3. Legal Implications

The theoretical section continues with an inquiry on the legal implications involved in using GIS data. Among the respondents, 14 claimed that they are aware of the legal implications involved in the use of GIS, while only one respondent answered to be not aware of any legal implications. Among the legal concerns, the most often mentioned were:

- Data anonymity, confidentiality and privacy issues.
- Copyright and property rights implied in spatial data can restrain the sharing and reuse of this data.
- Raw data cannot be published.

To conclude, spatial data is sensitive. It needs to be collected and shared under voluntarily agreed conditions between the data collector and the respondent, and its sharing is often restricted to small scale due to possible conflict of interests (e.g., revealing the geographic location of pig farms with ASF outbreaks).

5.3. Technical Section

This section inquired on the technical details regarding the data flow, format of the spatial data, and tools used to conduct spatial analysis using GIS.

5.3.1. Metadata

The survey identified several GIS datasets that are currently used by EFSA colleagues. This refers to the GIS data only, not the attribute data that is combined with GIS data. More than 60% of the data comes from the following topics (Table 13)**Error! Reference source not found.**:

- Agriculture
- Climate
- Animal and plant health
- Land Use

It has been noted that environmental data is used by plant health as much as by animal health.

Table 13: Fields using GIS data at EFSA

Fields	Number of responses	% of the total responses
Agriculture	9	17
Climate	9	17
Animal and Plant Health	9	17
Land Use	7	13
Population	6	11
Epidemiology	5	9
Chemical Substances / Biological Hazards	2	4
Zoonoses	2	4
Other	4	8
Food consumption / Exposure	0	0
Pesticide, Veterinary, Drug Residues	0	0

5.3.2. Data providers and databases

Most GIS data at use in EFSA is provided by European Member States (MS), national authorities. Besides, data is retrieved from agencies that provide public databases such as Eurostat, ECDC and FAO.

National, supra-national institutions, and scientific publications:

- World Organisation for Animal Health (OIE)
- Global Climate Data (WorldClim)
- Food and Agriculture Organization of the United Nations (FAO)
- National Aeronautics and Space Administration (NASA)

EU Member States and EU bodies and agencies:

- European Commission Joint Research Centre (JRC)
- European Commission European Statistical Office (Eurostat)
- European Commission Soil Data Centre (ESDAC)
- European Commission Animal Diseases Information System (ADIS)
- European Centre for Medium-Range Weather Forecasts (ECMWF)
- European Centre for Disease Prevention and Control (ECDC)
- European Space Agency (ESA)
- European Environmental Agency (EEA)

In general, it has been noted that GIS users retrieve spatial data mostly from public databases, and that there is not a single, centralized access point nor centralized repository for GIS data at EFSA. Although one respondent answered to retrieve GIS data from EFSA's Scientific Datawarehouse (DWH), most respondents claimed that there are still no standardized methods for GIS data access and retrieval at EFSA. (

- National databases from the authorities from Member States (MS)
- European Food Safety Authority (EFSA)

Table 14).

Field	Sub-fields	Database	Datasets
Animal Health	African Swine Fever (ASF)	Open source databases (e.g. data.europe.eu); Map / data services (e.g. ECDC GeoPortal); EFSA's Scientific Datawarehouse (DWH), including geographic information; Local database; national databases and statistics	NUTS; specialised shapefiles (e grounds) provided by M States; Climate (Temperature d mean and max); ASF ar animal disease data (PP etc.)
	Avian Influenza (AI)	-	-
	Description of disease profiles	Corine Landcover	Several ASF datasets
	Animal ecology	-	-
	Bee Health	EFSA's Scientific Datawarehouse (DWH) including geographic information; Local database; Open source databases (e.g. data.europa.eu)	NUTS
Zoonoses and AMR	Zoonoses	EFSA's Scientific Datawarehouse (DWH) including geographic information	zoonoses monitoring da
	Antimicrobial resistance (AMR)	EFSA's Scientific Datawarehouse (DWH) including geographic information	-
Plant health	pest risk assessment: risk of entry, climate suitability, crops/pests distribution	Open source databases (Ex. data.europa.eu); Paper-used databases ; Map / data services (ex. ECDC GeoPortal); Local database ; European Soil Database; Corine Landcover; WorldCover; Datasets used in scientific articles; Global Land Cover (Copernicus),	14 bioclimatic variables WorldClim, derived crop (Eurostat), Digital Eleva Models, Vegetation map LAND
Other	Sustainable food production	-	-
	Genetically Modified Organism (GMO)	-	meteorological data, soi plants, crop growing da management, data, lan mask.

The survey results identify the main databases and sources from which GIS users collect the input data. Some of the most commonly used databases are:

- The World Animal Health Information System (WAHIS), OIE
- Global Climate Data (WorldClim)
- EMPRES Global Animal Disease Information System FAO (Empres-I), FAO
- European Commission Soil Data Centre (ESDAC), EC
- European Commission Animal Diseases Information System (ADIS), EC
- European Centre for Medium-Range Weather Forecasts (ECMWF) - ERA5-LAND
- WorldCover, ESA
- Coordination of Information on the Environment (Corine Landcover), EEA

Field	Sub-fields	Database	Datasets	Source
Animal Health	African Swine Fever (ASF)	Open source databases (e.g. data.europe.eu); Map / data services (e.g. ECDC GeoPortal); EFSA's Scientific Datawarehouse (DWH), including geographic information; Local database; national databases and statistics	NUTS; specialised shapefiles (e.g. hunting grounds) provided by Member States; Climate (Temperature data, min, mean and max); ASF and other animal disease data (PPR, LSD, etc.)	EUROSTAT Member States (MS) EFSA's Scientific Data warehouse (DWH) SIGMA
	Avian Influenza (AI)	-	-	ADIS, OIE-WAHIS, Empres-I, Member States (MS)
	Description of disease profiles	Corine Landcover	Several ASF datasets	SIGMA, ADIS, Copernicus, EEA
	Animal ecology	-	-	Citizen science
	Bee Health	EFSA's Scientific Datawarehouse (DWH) including geographic information; Local database; Open source databases (e.g. data.europa.eu)	NUTS	EUROSTAT EUBEE
Zoonoses and AMR	Zoonoses	EFSA's Scientific Datawarehouse (DWH) including geographic information	zoonoses monitoring data	EFSA's Scientific Datawarehouse (DWH)
	Antimicrobial resistance (AMR)	EFSA's Scientific Datawarehouse (DWH) including geographic information	-	MicroStrategy, EFSA's Scientific Datawarehouse (DWH)
Plant health	pest risk assessment: risk of entry, climate suitability, crops/pests distribution	Open source databases (Ex. data.europa.eu); Paper-used databases ; Map / data services (ex. ECDC GeoPortal); Local database ; European Soil Database; Corine Landcover; WorldCover; Datasets used in scientific articles; Global Land Cover (Copernicus),	14 bioclimatic variables from WorldClim, derived crop maps (Eurostat), Digital Elevation Models, Vegetation maps, ERA5-LAND	Eurostat, Copernicus Land Monitoring Service, JRC, ESA, Nasa, ADIS, WorldClim, EEA
Other	Sustainable food production	-	-	Public databases EFSA's Scientific Datawarehouse (DWH) JRC
	Genetically Modified Organism (GMO)	-	meteorological data, soil data, host plants, crop growing data, crop management, data, land use, crop mask.	Public databases EFSA's Scientific Datawarehouse (DWH)

Table 14: Datasets used at EFSA

–: no answer. Some respondents' answers have been modified due to lack of consistency in the answers. The response "EFSA's geodatabase" has been here changed to "EFSA's Scientific Data warehouse (DWH) including geographic information". This table captures the dataset at use by field, with its corresponding data.
ADIS: European Commission Animal Diseases Information System; Corine Landcover: Coordination of Information on the Environment; DWH: EFSA's Scientific Datawarehouse (DWH); ECMWF: European Centre for Medium-Range Weather Forecasts; EEA: European Environmental Agency; EMPRES-I: Global Animal Disease Information System by FAO; ERA5-LAND: Climate dataset by ECMWF, Copernicus: European Union's Earth observation programme; ESA: European Space Agency; EUBEE: Partnership Prototype Platform on Bee Health; EUROSTAT: European Statistical Office; JRC: Joint Research Centre; LSD: Lumpy skin disease; NUTS: Nomenclature of Territorial Units for Statistics; OIE-WAHIS: World Organisation for Animal Health (OIE) World Animal Health; PPR: Peste des Petits Ruminants; SIGMA: EFSA project on Animal Disease Data Model; WorldClim: Global Climate Data; WorldCover: ESA worldwide land cover mapping.

5.3.3. Data Flow - Data harmonization and shareability

The results of the survey indicate that there is still not a centralized flow of GIS data at EFSA (Figure 14). Also, the survey responses show that in general there is little use of a pre-designed workflow for data extraction and standardized data models (20% use predesigned workflow; Figure 15). A common approach to access and use GIS data is to find and download the data of interest from an external database and import the GIS data in the activity or project that is being conducted.

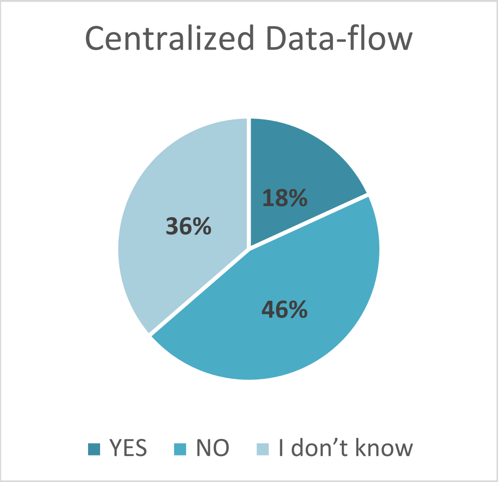


Figure 14: Centralized data-flow at EFSA

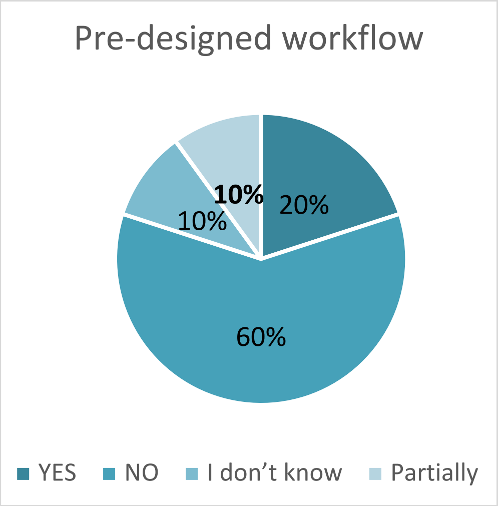


Figure 15: Pre-designed workflow for data extraction and standardized data models at EFSA

Further, the results show that most of the datasets are partly harmonized (73%; Figure 16) shows that interoperability is restrained since according to the survey results 9% of the datasets are not harmonized in terms of controlled terminology, geographic coordinates, projection, etc. This implies that 1 out of 11 outputs might not be comparable to others.

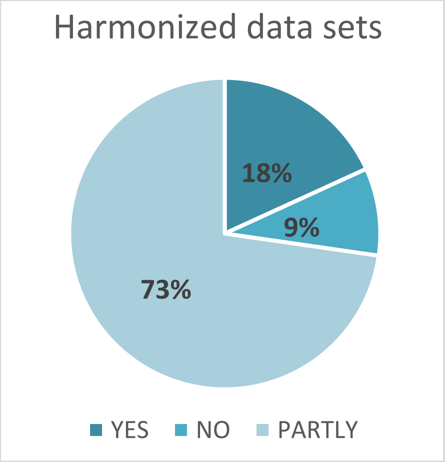


Figure 16: Harmonization of datasets at EFSA

5.3.4. Storage location of input GIS Data

The results of the survey also show the storage location of the input GIS data. It has been identified that there is currently no centralized repository for GIS data as more than 60% of the data is stored in individual, local files (Table 15).

Table 15: Location of input GIS data

Location	Percentage (%)
Spatial files in ad hoc / opinion specific folders	36
Spatial files in a structured centralized folder system (Ex. OneDrive)	24
EFSA Scientific Datawarehouse (DWH)	16
Other	12
I don't know	8
Access Geodatabase (database management system to store and manage data)	4

5.3.5. Format, scale and type of GIS data

The following graphs present the format (Figure 17), scale (Figure 18) and type (Figure 19) of the GIS data at use. CSV is the most used format of input data followed by shapefile (Figure 17). The most used scales for input data are country and NUTS (Figure 18). 61% of input data is in vector type (Figure 19).

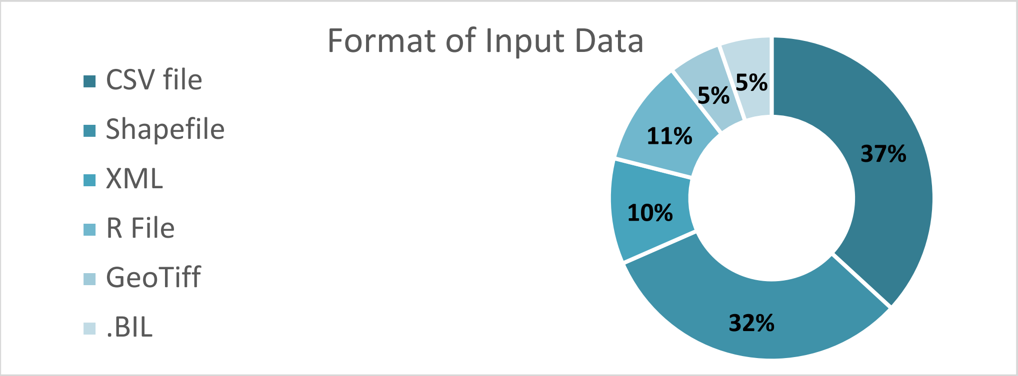


Figure 17: Format of input data

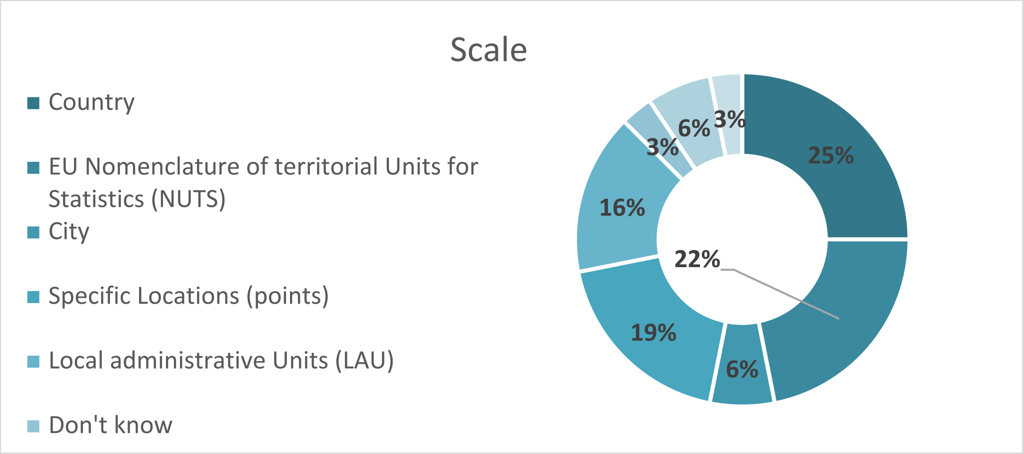


Figure 18: Scale of input data

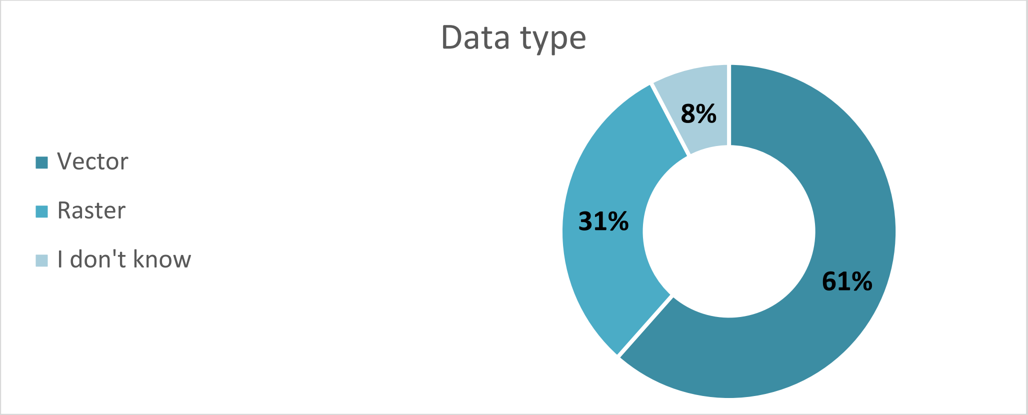


Figure 19: Type of input data

5.3.6. Elaborated spatial data: accessibility

The technical section continues to inquire about the output GIS data and its degree of accessibility. Results show that:

- 54% of the elaborated data is partially available to the public (in cases where sensitive data could be traced due to the data scale, information is geo-masked or anonymised).
- 31% of the elaborated data is publicly available.
- 15% is not available to the public.

The survey has collected information about means by which the broader public can retrieve the elaborated GIS data. Options listed are the following:

- Scientific reports and the EFSA journal
- EFSA Scientific Datawarehouse (DWH)
- Other (e.g. geoportal, webpage)

One respondent claimed that there is currently no system in place to make elaborated data available to the broader public. Although still limited to the internal access, one respondent claimed that it has been proposed to create a shared folder in Teams or SharePoint for storing elaborated GIS data.

Other respondents mentioned two databases where GIS data produced at EFSA is publicly available:

- Global Biodiversity Information Facility (GBIF) [25]
- Zenodo[26]

Only one respondent claimed that the output GIS data is not publicly available.

The output data can be retrieved in different formats depending on the activity or project. Most generic format for elaborated GIS data retrieval is Esri shapefile (shp, dbf, shx, tiff), followed by R files. Other formats used are CSV, HTML, GBIF annotated, geojson and json, gml, KML, KMZ, OSM and XML (Figure 20).

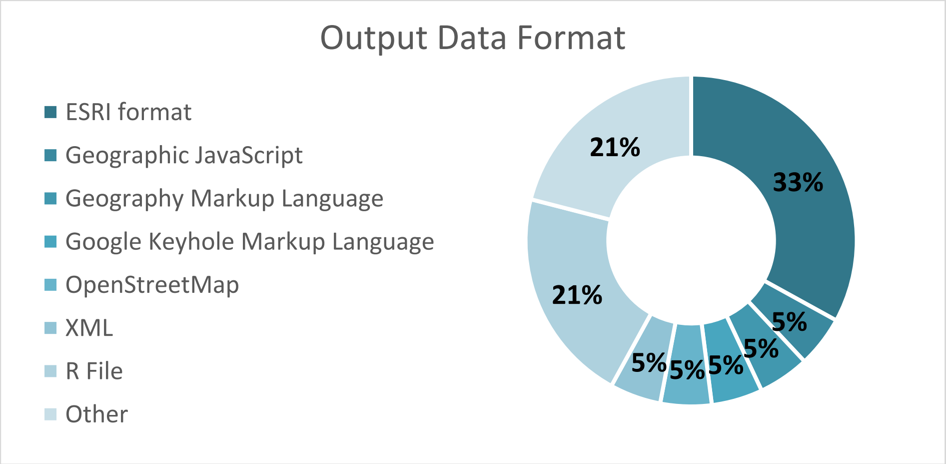


Figure 20: Format of output data at use at EFSA

5.3.7. GIS Tools

According to the results of the survey, different tools are used to conduct GIS at EFSA. More than 50% of GIS analyses are conducted using open-source tools such as R and QGis, as an alternative to ArcGis. Still, ArcGis continues to be used for conducting spatial analysis and data visualization in the form of story maps.

The following table lists the different tools that are used for GIS, differentiating between private and open-source software and presents user's justification for their preference (Table 16).

Table 16: Tools used at EFSA and users' justification for their preference

Tool	Percentage (%)	Accessibility	Justification
R	30	Open source	Process data in structured way
QGIS	18	Open source	Requested for specific purposes
ECDC EMMa	8	Proprietary software	Open access, simple, quick
ArcGIS	25	Proprietary software	Easy to use
Microstrategy	11	Proprietary software	Experience / knowledge
Access[27]	4	Proprietary software	Used in scientific community
R4EU	4	Proprietary software	Free

---

[25] https://www.gbif.org/es/

[26] https://zenodo.org/

[27] This option is presented as found in the survey responses. It is not clear if refers to Microsoft Access. Respondent clarified that Access is used for data manipulation.

Specific packages of R are used for conducting spatial analysis and visualization. The following list presents the packages mentioned by GIS users: leaflet, sp, rworldmap, rgdal, sf, raster, briskR, INLA, maptools, rgeos, tmap[28].

5.4. Descriptive Section

The descriptive section aimed at capturing the outcome material of any GIS activity, and identifying how it is used, and where it is published.

5.4.1. GIS outcome material

The survey identifies the most common GIS outcome material, how and in which context it is used. It is important to note that several answers were allowed for this section given that activities and projects can produce a combination of GIS outputs used in different contexts for varying purposes. The survey shows that the most common outcome material are static maps (31%) followed by internal interactive maps (Table 17). One respondent added story maps as an GIS outcome material adding that are useful tool to present and communicate a scientific output in a clear and interactive way.

Table 17: GIS outcome material

GIS outcome material	Percentage (%)
Static maps	31
Spatial dataset to be used in other analysis	14
Reports / tables	14
Interactive internal maps (accessible within EFSA only)	12
Graphs	12
Interactive web / external maps (publicly accessible)	10
Map / data services that external users can further re-use in their GIS analysis	5

Table 18 presents in detail the GIS outcome material by field.

Table 18: Detailed GIS outcome material by field

Fields	Sub-fields	GIS outcome material
Animal Health	Animal Health Law	Static maps
		Story maps
	African Swine Fever (ASF)	Static maps
		Interactive internal maps (only accessible within EFSA)
		Spatial dataset to be used in other analysis
		Interactive web / external maps (publicly accessible)
		Graphs
		Reports / Tables
	Avian Influenza (AI)	Static maps
	Description of disease profiles	Static maps
		Interactive internal maps (only accessible within EFSA)
		Interactive web / external maps (publicly accessible)
		Spatial dataset to be used in other analysis

---

[28] R packages at use at EFSA have been checked in R Cram to confirm that they are still functioning. Only Package BriskR was not found in the archive.

		Graphs
		Reports / Tables
		Map / data services that external users can further re-use in their GIS analysis
	Animal ecology	Static maps
		Spatial dataset to be used in other analysis
	Bee Health	Static maps
Zoonoses and AMR	Zoonoses	Static maps
	AMR	Static maps
Plant Health	Crops and pest distribution	Static maps
		Reports / tables
		Interactive web / external maps (publicly accessible)
	Plant pest and disease climate suitability	Static maps
		Spatial dataset to be used in other analysis
		Graphs
		Reports / tables
Other	GMO	Interactive web / external maps (publicly accessible)
		Spatial dataset to be used in other analysis
		Graphs

5.4.2. Use of GIS outcome material

The results of the survey capture how the GIS outcome material is being used. Notably, almost 50% of the GIS outcome material is used for visualization support, and 31% to support scientific analysis (Table 19).

Table 19: Use of GIS outcome

Use of GIS outcome material	Percentage (%)
Visualization support	44
Scientific support	31
EFSA scientific opinion	13
Environmental risk assessment of GMO	6
European Summary Reports (EUSR)	6

5.4.3. Context where GIS outcome material is published

Table 20 presents the context where GIS outcome material is published. More than 50% of the produced GIS material is presented in EFSA scientific Opinion and annual reports, while only 12% is shared on the web.

Table 20: Context where GIS material is published lister with percentages

Context where GIS material is published	Percentage (%)
EFSA scientific opinion	33
EFSA Annual Report	21
General EFSA Publications	12
Exposure and Risk Assessment	12
Web	12
EFSA Technical reports	3
Sustainability assessment	3
Applications hosted in R4EU	3

R4EU is platform hosting models used in EFSA assessments

5.5. Future GIS situation at EFSA

This section collected GIS users' experience and expectations to prepare for future implementation of GIS at EFSA.

5.5.1. Restrains encountered in GIS

The survey has collected several of the restraints that staff have faced while working with GIS (Table 21).

The restraints faced when using GIS come from the following reasons:

- Data quality issues (60%; Figure 21)
- Legal implications (25%; Figure 22)
- Tools and applications used to do GIS (16%; Figure 23)

Table 21: List of restraints encountered while using GIS

Restraint category	Description of restraints	Percentage (%)
Data quality issues	Spatial data is often missing or incomplete	25
Data quality issues	The description (documentation) of available spatial data is often incomplete	18
Legal implications	Cultural, institutional, financial and legal barriers prevent or delay the sharing and re-use of existing spatial data	15
Issues with tools and processes	Inadequate expertise on spatial analysis to conduct more in-depth analysis	13
Legal implications	Restriction on the scale of the data because of legal implications	10
Data quality issues	The systems to find, access and use spatial data often function in isolation only and are not compatible with each other	8
Data quality issues	Spatial datasets can often not be combined with other spatial datasets	8
Issues with tools and processes	Inappropriateness of the software or GIS tools	3
Issues with tools and processes	Complicated process needs to be followed to request the production of maps to another EFSA unit	3

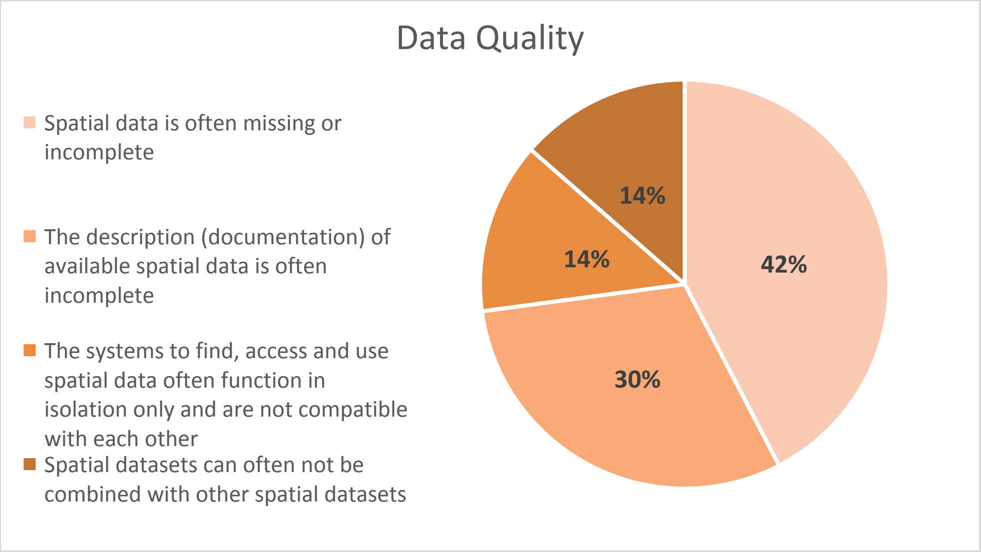


Figure 21: Data quality

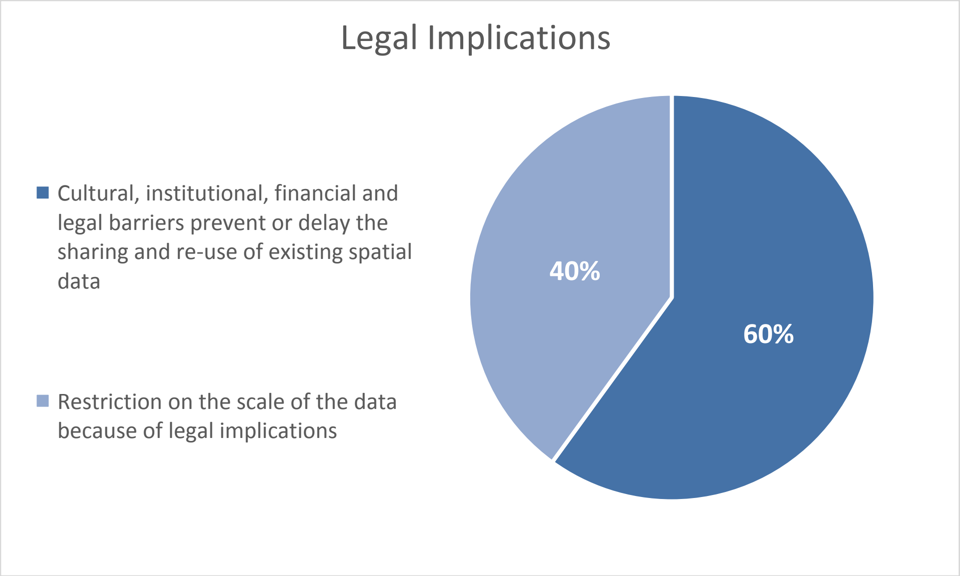


Figure 22: Legal implications

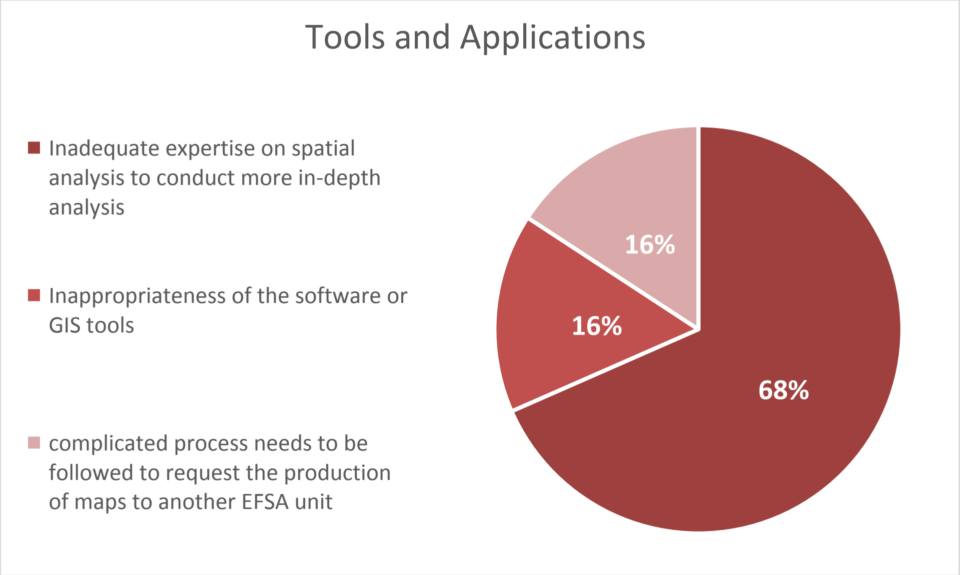
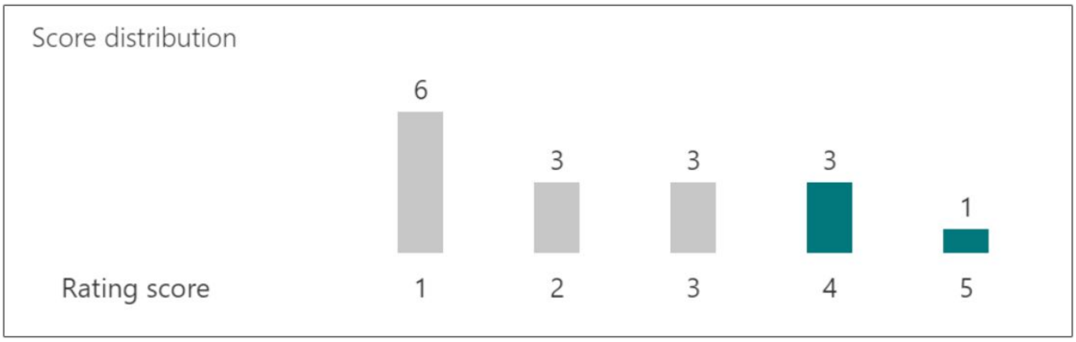


Figure 23: Tools and applications used to do GIS

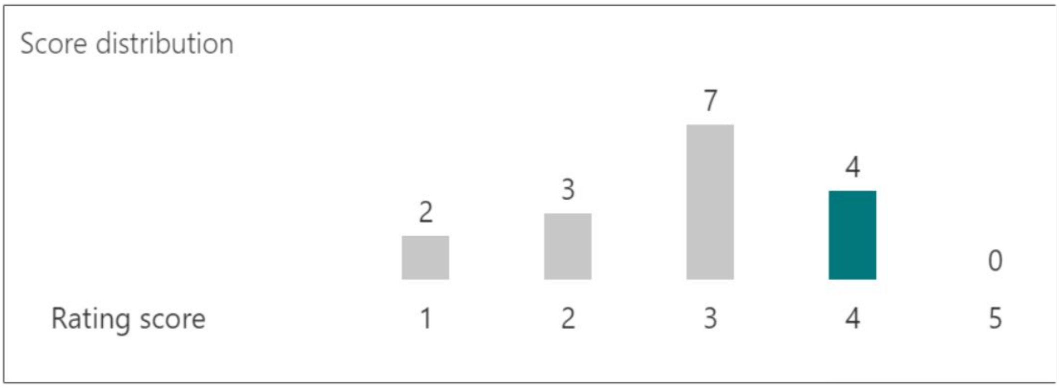
Another common limitation encountered among GIS users at EFSA is the technical complexity entailed in GIS. The survey results show that the level of formal education or experience in GIS is relatively low, with 40% of the respondents answering 1 on a scale from 1 to 5 (1 being the lowest; Figure 24). Still, 25% rated themselves to be between 4-5 level of formal education. Similarly, the level of expertise in using GIS tools and understanding the spatial data format is also low on average, with a median of 3 out of 5 (Table 22).



Level of formal education or experience in GIS, on a scale from 1-5, 5 being the highest

Figure 24: Formal education or experience using GIS

Regarding the level of expertise on using GIS tools and understanding the spatial data format, 25% rated between 4-5 in a scale from 1-5, 5 being the highest value (Figure 25)**Error! Reference source not found.**.



Level of expertise on using GIS tools and understanding the spatial data format, in a scale from 1-5, 5 being the highest.

Figure 25: Expertise using GIS tools

Considering the level of expertise (most frequent value of 3; Figure 25) in relation to the formal level of education (most frequent level of 1; Figure 24) leads to the conclusion that most of GIS users at EFSA have learned to use GIS during their working years, without receiving extra education on GIS. In fact, the survey results captured GIS users' interest and need for formal education on GIS provided by EFSA.

Table 22: Technical expertise limitations

Limitations in technical expertise	Min	Median	Max
Level of formal education or experience in GIS	1	2	5
Level of expertise on using GIS tools and understanding the spatial data format	1	3	4

Minimum, median and maximum of the level of formal technical education and level of expertise on a scale from 1 to 5, 1 being the lowest.

5.5.2. Recommendations for future GIS implementation

The survey results capture that the use of GIS at EFSA can possibly increase in the future. 45% of the respondents answered that are planning to conduct or participate in a future project or activity working with GIS. Only 10% are not using GIS in the future. The rest of respondents will possibly use GIS in future activities. Regarding GIS data, 50% of respondents answered that expect to receive GIS data in the future, while the other half is not expecting to receive it.

Given this possible increase of GIS at EFSA, the survey collected GIS users' recommendations for future GIS implementation based on their past experience. The results capture recurrent observations among respondents. Some of the common recommendations are:

- Well defined data models for **data collection**, for which the spatial component is well defined and spatial resolution is agreed on in order to ensure collection is harmonized between EU MS.
- More clear and harmonized structures for **data management**
- Need of **common modelling** framework able to use GIS data and re-use modelling components for different purpose under the remit of EFSA.
- Common **data architecture** and development of common modelling framework
- Foster **collaboration** among international agencies
- Need of **centralized** data repositories
- Promote **training** in GIS and remote sensing
- Increase spatial data **accessibility** to EFSA staff
- Make the collected data **open** and available
- Harmonize the type of platform for the interactive use of maps from outside users
- Evaluate the use of open source software
- Increase the number and type of software licenses available, including remote sensing applications

5.6. GIS at EFSA - Discussion

The results of the survey were shared with the respondents in a meeting held on November 2021.

The meeting consisted of the presentation of general conclusions on the current situation of GIS at EFSA based on the survey results. The meeting also opened a discussion on future GIS implementation at EFSA based on the answers collected in the survey.

Three key issues were raised by participants during the discussion. First, the lack of standards for data storage, use and analysis is not only a challenge inside EFSA. The situation outside EFSA faces similar challenges due to the copious amounts of data, technology limitations, etc. This challenge is to be tackled with the SEED project.

Second, it was discussed about the importance of raising awareness about the use of GIS as a future EFSA strategy, particularly under the existing needs to tackle the climate crisis from a holistic approach.

Third, it was mentioned that, besides its use for risk assessment, GIS is also a powerful tool for communication due to the possibility to visualize data and findings, for example with story maps.

The discussion offered a space to share experiences on past work with GIS at EFSA. As such, one EFSA colleague shared existing material on a previous initiative on the use of GIS at EFSA. This information has been collected in a shared folder as historical GIS material.

5.7. EFSA in collaboration with other EU bodies: GIS case

The previous section has presented the results of the activities and projects using GIS at EFSA as tracked through the survey. This section aims to present an example of a GIS platform that results from the collaboration between EFSA and other European bodies.

5.7.1. Information Platform for Chemical Monitoring (IPCHEM)

The Information Platform for Chemical Monitoring (IPCHEM), is the European Commission’s reference point to access chemical occurrence data collected in Europe. It is a centralized data platform that combines chemical monitoring data across disciplines and institutions. IPCHEM is a coordinated approach, between the European Commission, including DG Environment, DG Sante, and the JRC, European Agencies (EFSA and the EEA), Member States Bodies and Research Centres. Its aim and scope is collecting, storing, accessing and assessing data related to chemical occurrence and mixtures, in relation to humans and the environment.

IPCHEM objective is to “identify the links between exposure and epidemiological data in order to explore biological effects and lead to improved health outcomes” (IPCHEM, 2021)). For this purpose, IPCHEM collects data related to chemical monitoring from different areas, which are structured in four modules; Environmental Monitoring, Human Bio-monitoring, Food and Feed, and Products and Indoor Air. EFSA contributes to IPCHEM with the submission of chemical monitoring related to food and feed, for example data of pesticides residues in food and feed and occurrence of chemicals in food.

5.7.2. GIS role at IPCHEM

IPCHEM aims at supporting users in spatial data visualization and spatial analysis in chemical exposure assessment. For this purpose, IPCHEM has upgraded the platform to provide a spatial module that facilitates spatial data access and visualization.

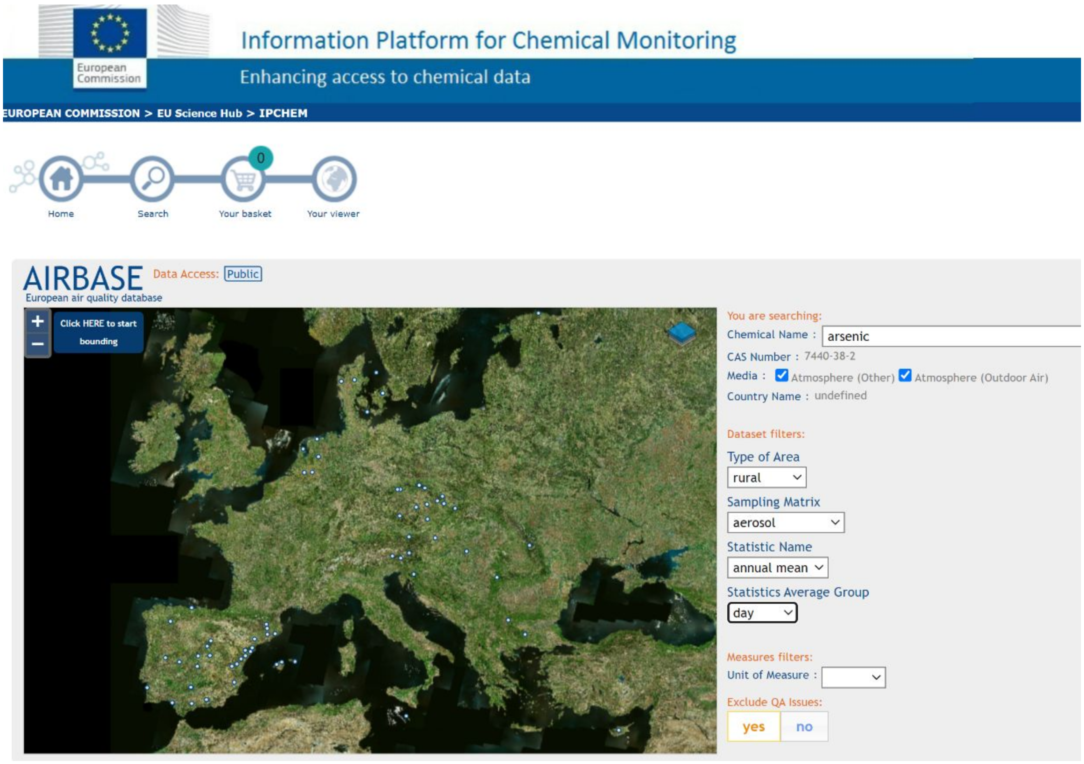
The application provides the further objectives:

- Facilitate spatial visualization and analysis of exposure assessments and in the interpretation of spatial and temporal trends of exposure sources
- Improve communication and support policy making as it facilitated exposure and risk assessment practices in support of EU policies

5.7.3. Geospatial infrastructure

The IPCHEM spatial module consist of a public-access GIS application that provides the user with an user-friendly graphical user-interface (GUI) to access, select, and visualize spatial chemical monitoring data related to human and environmental health, food and feed, and air data (Figure 26). The tool provides a menu to select data and apply filters in a query to choose media, project/institution, date, and geographic scale (Europe, world, and country level).

The application creates a dynamic map that visualizes the selected data on a basemap. It allows the user to interact with the map to select an observation and get specifications on the selected sample site.



You are searching:
Chemical Name : arsenic
CAS Number : 7440-38-2
Media : ☑ Atmosphere (Other) ☑ Atmosphere (Outdoor Air)
Country Name : undefined

Dataset filters:
Type of Area: rural
Sampling Matrix: aerosol
Statistic Name: annual mean
Statistics Average Group: day

Measures filters:
Unit of Measure :
Exclude QA Issues: yes no

Figure 26: Screenshot of the IPCHEM spatial module

5.7.4. Application advantages

The main advantage of the spatial module is the possibility to combine data across different disciplines, geographic boundaries and European Institutions. Hence, the IPCHEM spatial module combines chemical data related to human and environmental health, food and feed, and air data, which allows to entail the link between chemical monitoring and human and environmental health. Besides, the collective platform allows to combine and centralize data provided by EC Directives and JRC, European Agencies (EFSA, EEA), and other European institutions. The Platform allows to provide chemical monitoring data in terms of "spatial, temporal, methodological, and metrological traceability" (IPCHEM, 2021) and allows to entail the link between chemical monitoring and human and environmental health.

Key messages

- There is the need for centralized geographic data management processes and practices that allow spatial data interoperability and models harmonization.
- The use of GIS at EFSA can possibly increase. It is currently mostly used in the domain of biological monitoring.
- At EFSA, the use of GIS and spatial statistics applied for biological monitoring allows to transition from description to simulation. The main purpose of using GIS is to identify, describe and prognosticate the geo-distribution of and event (e.g. disease) and identify potential (geographical) risk factors.
- GIS users use private and open access tools to conduct GIS, but more than 50% of GIS analyses are conducted using open-source tools such as R and QGis
- GIS products, such as maps and story maps, are useful for improving visualization and communication of scientific outputs to policy makers.
- Regarding data collection, the majority of GIS data at use at EFSA is provided by Member States. At the same time, EFSA collects its data from other supra national bodies that offer a public databases such as FAO and the OIE.
- GIS capability is restrained by data quality issues, legal implications, and tools and applications used to do GIS
- Recommendations point out towards harmonized practices for data collection, management and modelling, as well as the need to increase data accessibility and centralization. More training in GIS was mentioned by four respondents.
- IPCHEM is a good example of a collaborative project among different public bodies with a spatial component that allows to combine data across disciplines, geographic boundaries and institutions.

6. Case Study

The findings of the scoping review, the visual exploration and the tracking of GIS activities at EFSA have identified the advantages and need of GIS web applications to facilitate spatial data access and use. The scoping review has identified the rise of a Web GIS due to the move towards web-based applications, the use of cloud computing and open source tools, while the visual exploration has helped to identify various web applications (geoportals and geodatabases) offered by several agencies to easily access, use, and visualize spatial data. Lastly, tracking the current state of GIS at EFSA has identified that more than 50% of GIS users at EFSA use open-access/ open-source tools to conduct spatial analysis (QGis, ECDC EMMa, R). It also highlighted the need of standardized spatial data management and GIS self-served tools to easily access and visualize spatial data. Therefore, given the identified widespread practice and benefits of geoportals and self-served GIS services, we have developed a GIS application using open-source tools as a case study. The case study had the aim of exploring the viability of developing a GIS application using open source solutions (Python) and testing the feasibility of providing self-served GIS services to access and use spatial data.

6.1. Description of the tool

A cross-platform application has been developed using Python as a programming language to provide GIS self-served services. Python is a high-level programming language that is powerful, flexible and easy to use and read. It comes under the Open Systems Interconnection (OSI) approved open-source license which makes it free to use and distribute. One of its most important advantages is that it can count on a vast library support and a large active community base.

The tool (Figure 27) provides the user with a user-friendly graphical user interface (GUI) to access, select and visualize EFSA spatial data in interactive maps. The maps consist of three layers:

- Basemap;
- Labels;
- Data.

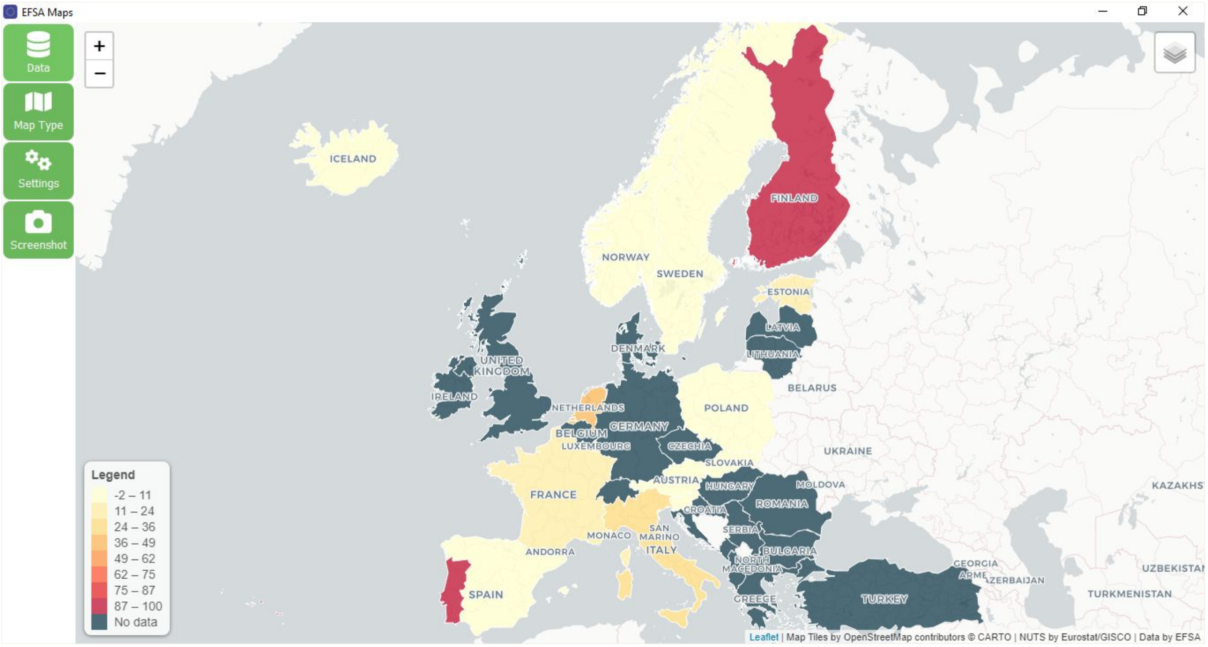
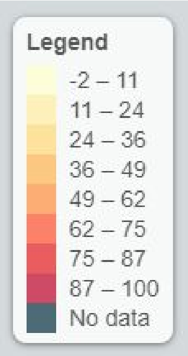


Figure 27: Screenshot of the tool

The basemap provides the overall visual context for the map and contains geographic features such as regions, or administrative boundaries.

The labels layer provides the text positioned on the map in relation to map features (such as country or city names, etc.).

Finally, the data layer is the most crucial one as it displays the geographic data on top of the basemap layer. The data source for the layer is a csv file which is retrieved from the EFSA systems. This is then converted into a GeoJSON file containing the territorial units for statistics (NUTS) established by Eurostat. The data is shown as a choropleth map, which enables to easily visualize a variable's variation across a geographic area. Maps are also provided with a legend as shown in Figure 28, explaining the symbols used to represent the geographic data on the map as well as a description of the corresponding range of values.



Legend

- -2 – 11
- 11 – 24
- 24 – 36
- 36 – 49
- 49 – 62
- 62 – 75
- 75 – 87
- 87 – 100
- No data

Figure 28: Map legend

All these three layers can be toggled through a button positioned at the top right corner (Figure 29) and can be shown or hidden according to the user's needs.

☑ Basemap
☑ Labels
☑ Data Layer

Figure 29: Layers can be toggled through a button

The GUI was developed using the PySide6 module which is a Python binding of the cross-platform GUI toolkit Qt. This library is available under GNU Lesser General Public License (LGPL) published by the Free Software Foundation (FSF) and allows developers to distribute their application without the need to release the source code.

The interactive maps are implemented using the pyqtlet repository, a Leaflet map wrapper for Qt bindings. Leaflet is a modern open source JavaScript library for developing interactive maps and focuses on performance, usability, simple API, small size and mobile support.

Unfortunately, at the time of writing this report, pyqtlet is no longer in active development and many Leaflet features remain unimplemented. However, a positive side is it can run JavaScript code and allows developers to access all the features available in Leaflet even though those are not implemented in pyqtlet. The map implementation code therefore contains both Python and JavaScript code.

6.2. Lessons learnt

Several advantages were found when developing the application. First of all, it was relatively simple to develop an interactive map application using Python as well as modern open source tools. Python can be considered a beginner-friendly programming language, which allows to easily develop a basic tool that permits to visualize spatial data through an interactive map.

Another advantage of such a tool is its user-friendly interface that allows to select and plot data in an interactive, visually appealing and straightforward map. It also allows to directly access data through APIs, procuring a centralized flow of spatial data and respecting privacy policies subjected to the data.

This application can also be seen as a self-served tool that allows to automatize the elaboration of maps and eases the access and use of spatial data. Therefore, it does not require any GIS technical skills, thus helping to overcome technical barriers to the use of spatial data, encouraging the use of GIS.

An open source software is free to use and could minimize organization's expenses by saving on licensing fees. On the other side, open source libraries might also imply challenges such as security, compatibility and maintenance. In our specific case, we have seen how the pyqtlet library is no longer maintained which could represent a problem for an organization's needs.

7. Recommendations for future GIS at EFSA

1. The strength of GIS lays in its capacity to combine data across different disciplines and geographic boundaries. Under EFSA's aim to endorse the One Health concept, GIS should be promoted as a tool that increases the collaboration across other European agencies and offers a holistic approach to food safety.

2. At EFSA, GIS is a useful tool for improving visualization and communication of scientific information to policy makers by presenting complex findings in straightforward maps. Hence, specially under EFSA's attempt to improve visualization and communication products, GIS material, such as story maps and dynamic maps, should be fostered.

3. GIS provides the platform to conduct spatial analysis that allows to transition from description to simulation. The potential of GIS should be further explored and promoted across EFSA's remit, with the aim to move from mostly data visualization to strengthen statistic spatial analysis.

4. There is currently no single, centralized access nor storage point for GIS data at EFSA. There are still no standardized methods for GIS data retrieval at EFSA. However, 25% of GIS users claimed to collect spatial data already from EFSA's Scientific Datawarehouse (DWH). Still, a centralized GIS data repository is needed in order to support the different activities through standardized approaches.

5. There is currently no centralized flow of GIS data at EFSA. In general, there are no standardized models for data management and analysis for GIS data, whereas for other domains standardized data models, strong data collection and analysis framework are implemented (e.g. SSD2 and GDE2). At the same time, spatial data models are not fully harmonized, which constrains data integration, interoperability, and replicability of results. There are currently several initiatives at EFSA to procure future harmonized and standardized practices for spatial data collection, management and analysis of environmental spatial explicit data (e.g. SEED).

6. A common limitation encountered among GIS users at EFSA has been the technical complexity entailed in GIS. Even though the level of formal education on GIS among EFSA's users is low, the level of expertise is satisfactory, suggesting that a minimum of education would sensibly increase GIS users' expertise and could possibly maximize its potential.

7. The potential of GIS at EFSA can still be maximized, as GIS activities are still scattered, and the survey results corroborated that GIS users are not identified within a GIS community. This is partly due because GIS activities and projects lack visibility at EFSA. Increasing GIS' visibility can possibly encourage its use and enforce the centralization of GIS data management and services, which can maximize the resources' impacts (tools, data, knowledge) and allow an efficient response to GIS needs.

8. According to the survey results, more than 50% of GIS analyses are conducted using open source tools such as R and QGis. Most of the elaborated spatial data is accessible to the public as maps or story maps published in EFSA's scientific reports and the EFSA journal. GIS output and services should be increasingly accessible to the public, while respecting data protection policies. Given the advantages of a GIS web-infrastructure and user-friendly interfaces, EFSA should consider moving towards a GIS web infrastructure, e.g. a user-friendly geoportal providing self-served services that facilitate the internal and public access and visualization of data relevant to EFSA's remit. This move would follow a body of policies on public data shareability, accessibility and transparency, as seen in the INSPIRE Directive, and is based on other agencies' initiatives to promote GIS data accessibility (see UWWTD-SIIF: winner of the geospatial policy implementation excellence award 2017 for its open source platform to monitor urban waste water treatment). An application has been developed and included in this technical report as a proof of concept to test the feasibility of providing self-served GIS services using open-source solutions. This case study could be used as a preliminary proof of concept for future analysis on the implementation of a GIS application and spatial services.

9. Future implementation of GIS should give priority to the domains that have been identified as the most impacted by GIS: biological monitoring (e.g. animal and plant health), data management, and analysis activities, to then explore the potential of GIS across the EFSA's remit.

10. Users should be fully aware of the legal implications regarding spatial data access and shareability. The use and sharing of spatial explicit data is limited by legal implications such as data anonymity, confidentiality and privacy issues.

8. Discussion and conclusions

The current GIS scenario has been explored in academia, through the scoping review, and in agencies and public bodies, through the visual exploration. Specific attention has been given to the use of GIS in the field of food safety, by exploring EFSA's GIS scenario.

GIS has proved to be a versatile tool that is applied in numerous diverse fields (spatial epidemiology, ecology, agriculture and soil science). It has been identified that EFSA uses GIS mostly for biological monitoring, with a focus on animal and plant health. The increase of GIS usage in these fields is due to the rising awareness of the geographical place and temporal dimension at which health events emerge and disperse. It has been highlighted throughout the report that GIS' advantage is its ability to combine data across disciplines. However, for the purpose of organizing the findings of the scoping review, these were grouped by field (public health, epidemiology, ecology, and agriculture and soil science). Nevertheless, the report acknowledges that, in line with the One Health concept, health (human, animal and plant), environment, and society is interconnected, and GIS provides the capacity to analyse these with a holistic approach by combining data across different disciplines and geographic borders. This holistic approach to data analysis also overcomes data siloes, as all data can be integrated through the spatial and time components. Another advantage of GIS is the possibility to combine attribute with spatial data, illuminating spatial patterns that might pass unseen.

Further, it is important to acknowledge the limitations encountered throughout the elaboration of this report. Regarding data collection and analysis, the scoping review was limited because the literature was restricted to reviews, which implied working with pre-distilled information. This limited the detail and quality of data as it was not directly taken from a primary source. Yet, the decision of working with reviews was suitable for this report, as we were not interested in specific outcomes of studies using GIS, but rather on an assessment of the uses of GIS in diverse fields. Further, given the diversity of data collected through reviews, and the exploratory nature of the report, it was a challenge to design a form to consistently extract the data. This difficulty was solved by combining open-ended with close-ended questions, and conducting a qualitative and quantitative analysis. Besides, another challenge encountered given the quality of the data from reviews was classifying the methods used to conduct spatial analysis. Methods were discussed by reviews at a different level of detail, and no standardization in terms of their application was found across the reviews. Therefore, we decided to capture the general practices of spatial analysis, rather than providing a specific list of methods. Moreover, the visual exploration was limited in terms of the number of agencies covered and the scope of the findings retrieved by visually exploring their websites. However, it was a good exercise to collect examples of applications of GIS services to facilitate spatial data access and visualisation, as exemplified by the identified geoportals and geodatabases. Lastly, regarding the limitations encountered in this report, we would like to note that the scope of this analysis was geographically limited to North America and Europe, because of the relevance to EFSA. Further research should examine how GIS is used in the domains of public health and food safety in other regions beyond the ones explored in this report.

This report has explored the GIS scenario in terms of the fields of application of GIS, and the data, infrastructure, and methodologies used to conduct spatial analysis. It has noted that data is collected at faster and larger quantities than ever before thanks to the development of collection and processing tools. Also, it revealed that there is a joint effort towards data accessibility and shareability. This is confirmed in scientific literature, as the scoping review findings show a higher mention of public datasets. It goes in hand with a high choice for open data, as seen in geodatabases, and self-services provided by geoportals to facilitate spatial data access and use, as seen in the agencies and public bodies explored. At the same time, data is collected at an unprecedent fine scale, which allows to expand the study of health across a time-space continuum. However, these immense amounts of data collected at increasingly higher spatial and temporal resolutions at higher frequency pose challenges regarding the use, share and analysis of GIS data. Besides, there is a trade-off between data resolution and cost. Processing and analysing copious amounts of data implies more complex models, difficult to undertake and interpret, and demands high hardware storage and computational capacities. Finer scale implies more policies related to geoprivacy, which limits data shareability and publication. Indeed, GIS users at EFSA acknowledge that GIS capability may be restrained by legal implications, that could even prevent the collection of spatial data at finer resolution. Besides, it is important to highlight that standardized practices to process data are not fully put in place, nor are standardized models to analyse it, and that spatial data is still limited by data fragmentation and siloes, because of competing interests and variety of formats. The scoping review and the visual exploration highlight these challenges in the GIS community, and they are also faced internally at EFSA. Indeed, it is EFSA's goal to overcome these challenges and achieve harmonized GIS data practices.

Regarding the infrastructure to conduct GIS analysis, exploring agencies and public bodies has revealed a current presence of geoportals and geodatabases that provide self-served services to access and visualize data. The fundamentals behind these applications lie in an increasing push towards open access to data and are motivated by a body of policy regarding public data shareability, data accessibility and transparency. However, data access provided by geoportals and geodatabases is also subjected to policies regarding geoprivacy, especially when scale is fine, and human subjects are involved. The main advantage of geoportals is the easy creation of (interactive) maps, which allow to visualize and communicate scientific findings, overcoming previous technical barriers to access and use geospatial data. To facilitate a centralized data flow, the geodatabases can be connected to geoportals, as seen in the ECDC Geocatalogue and the map maker tool EMMa. Besides, these interfaces aim at translating data into easily accessible, valuable information for a range of end-users, from researchers to policy makers. Also, centralized geodatabases are key to an efficient retrieval of accurate up-to-date data. The vast accessibility of data now creates a challenge: the many options of portals and points to access data complicates the retrieval of information. Part of data connectivity entails joining data repositories and having single data access points (Saligoe-Simmel in EU Esri Summit, 2021). At EFSA, GIS activities are not fully connected, and practices to collect, process and use spatial data, as well as tools to conduct spatial analysis are not fully standardized.

Also, the importance of metadata has been highlighted. Well-documented metadata is crucial for facilitating data retrieval and reusability. The addition of metadata for discovery and measurement methods allows combining individual data from different fields of interest, as recommended by the USDA Global Agricultural & Disaster Assessment System (GADAS). Examples of well documented datasets are: INSPIRE Geoportal, European Data Portal, ECDC Geo Catalogue, and the EEA Geospatial Geo-catalogue.

Finally, the scoping review identifies a migration of services to the web, which gives rise to a Web GIS organizing principle, which allows to integrate databases, collect data, and disseminate results thanks to the internet and inexpensive storage capacities. This helps overcoming data siloes. However, it is important to highlight that storing data remotely may lead to legal issues regarding confidentiality. Similarly, the case study served to point out the challenges encountered when relying on open-source tools to provide GIS services. Despite the application developed for this case study proving to be successful in easily developing an interactive map application using open-source solutions, these might also lead to challenges such as security, compatibility and maintenance.

Regarding methods, GIS provides the medium to conduct spatial statistical methodologies that allow to move beyond description to simulation, which supports the transition of focus from control to prevention of events. Besides, other benefits of spatial analysis are facilitating data exploration and interpretation. Spatial methods are used interchangeably for different analyses concerning public health. Currently there is an exponential growth of epidemiological applications of Bayesian spatial modelling given the advances in computational power. At EFSA, complex spatial analyses using spatial statistics are used for animal health monitoring (ASF), plant health (plant/pests risk assessment and climate suitability), and modelling disease profiles (vector-borne diseases).

We have presented the different challenges that GIS currently faces, such as harmonizing spatial data across geographic boundaries and disciplines and easing data access and interpretation. However, we conclude that although GIS faces challenges, it is a powerful tool which is increasingly being used in the domains of public health and food safety to study the relationship between health, space and time. We argue that one of the main benefits of GIS is the possibility to approach health from a holistic view, which combines information across fields and geographic boundaries by a common component: space. Specially under the current climate and sanitarian crisis, it is urgent to approach health holistically, as recognized by the One Health Concept, under which the environment, society and health are intrinsically interconnected. At the same time, attention in the domains of public health and food safety is transitioning from control to prevention, and GIS offers the technology to conduct spatial analyses that allow to move from description to prediction. Lastly, GIS has proven to be an advantageous communication tool as it allows to present complex scientific findings in a straightforward and dynamic way, such as in maps or story maps. Finally, there has been already a remarkable success in transitioning towards better spatial data accessibility and increasing public access to geographic data, seen in initiatives such as the European Inspire Directive, and the European Data Portal. Yet, more has to be done in order to achieve complete data harmonization, interoperability, accessibility and integration in order to support risk assessment that leads to informed decision making.

9. Appendix

Appendix I - Complete list of all agencies and supra-national bodies exploredTable 23: Complete list of all agencies, and supra-national bodies explored. Area, Agency, Field, Name of application, Type of application, tool used to provide geospatial service, distribution, licencing and tool accessibility.

Area	Agency	Field	Name	Type of Application	Tool	Services	Distribution	Licencing	Access
Chemical monitoring	EC	Chemical monitorin, human and environmental health	IPCHEM	Portal with spatial component			Web-based	Proprietary (Creative Commons Attribution 4.0 International (CC BY 4.0) licence)	Free
Animal and Plant Health	USDA	Animal and Plant Health	APHIS Geoportal	Geographic Portal software services (Geo Portal)	ESRI		Web-based	Proprietary	Restricted
Human Epidemiology	CDC	Human epidemiology: HIV, viral hepatitis, STDs, and TB	Atlas Plus	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free
ALL	JRC	European Data: Land Use, Land Cover, Soil, Copernicus, Population, Environment	Big Data Analytics Platform (BDAP)	Geo Database	-	Download	Web-based	Proprietary	Free
Human Epidemiology	CDC	Human epidemiology	chronic disease indicators (CDI)	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free
ALL	EC	European data: Economy and Finance / Education, Culture & sport / Energy / Environment / Government and Public sector / Health / International Issues	Data Europa	Geo Database	-	Download	Web-based	Proprietary	Free

		/ Justice, Legal System & Public Safety / Population and Society / Regions and Cities / Science and Technology / Transport / Agriculture, Fisheries, Forestry and Foods							
Public Health	CDC	Public Health: Disability	Disability and Health Data System (DHDS)	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free
Environment	EEA	Environment (land, climate, bio, copernicus, marine, image, air, water, SOER, LDP, Noise, maratlas, citizen science))	Disco Map	Geo Database	-	Download	Web-based	Proprietary	Free
Human Epidemiology	CDC	human epidemiology - Bacterial Diseases	Division of Bacterial Diseases - NCCDPHP		-				NA
Human Epidemiology	CDC	Global Migration and Quarantine	Division of Global Migration and Quarantine		-				NA
Public Health	CDC	Population health	Division of Population Health - NCCDPHP		-				NA
Public Health	CDC	Reproductive Health	Division of Reproductive Health - NCCDPHP CVC	Geo Database	-		Web-based	Proprietary	Free
Human Epidemiology	ECDC	Disease Prevention and Control	ECDC EMMA	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free
Human Epidemiology	ECDC	Disease Prevention and Control	ECDC Geo Portal	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free
Human Epidemiology	ECDC	Disease Prevention and Control	ECDC Geocatalogue	Geo Database	-	Download	Web-based	Proprietary	Free

Environment	EEA	Environmental data (in line with INSPIRE topics): Agriculture, Environment and Health, Land Use, Soil)	EEA Geospatial Data Catalogue	Geo Database	-	Download	Web-based	Proprietary	Free
Public Health	CDC	Public health and Epidemiology	Epi Info	Statistical analysis software	-		Local	Proprietary	Free
Food and Drug	FDA	Food and Drug Administration	FDA GeoPortal	Geographic Portal software services (Geo Portal)	-		Web-based	Proprietary	Restricted
Agriculture	USDA	Global Agricultural & Disaster Assessment System	GADAS Global Agricultural & Disaster Assessment System	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free
Human Epidemiology	NCI	Human Epidemiology: Cancer	GIS portal for cancer research	Geographic Portal software services (Geo Portal)	ESRI	Download	Web-based	Proprietary	Free
Statistical Units	Eurostat	Administrative boundaries (NUTS) for statistical maps. Geospatial information for the Commission	GISCO	Geo Database	-	Download	Web-based	Proprietary	Free
Public Health	CDC	public health, human epidemiology	GRASP		-				NA
Statistical Units	EC / Eurostat	Professional statistical maps with Eurostat Data	IMAGE	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Restricted
ALL	EC	Spatial Information in the European Community: Annex I, II, III	INSPIRE Geoportal	Geo Database	-	Download	Web-based	Proprietary	Free
Human Epidemiology	CDC	Human epidemiology - Division for Heart Disease and Stroke Prevention	Interactive Atlas of Heart Disease and Stroke	Geographic Portal software services (Geo Portal)	ESRI	Download	Web-based	Proprietary	Free
Human Epidemiology	CDC	Human epidemiology: Heart Disease and Strokes	Interactive Atlas of Heart Disease and Strokes CDC	Geographic Portal software	ESRI	Download	Web-based	Proprietary	Free

				services (Geo Portal)					
Human Epidemiology	CDC	Human epidemiology - diabetes	Interactive Diabetes Atlas	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free
Public Health	CDC	Nutrition, Physical Activity, and Obesity	Nutrition, Physical Activity, and Obesity: Data, Trends and Maps	Geo Database	-		Web-based	Proprietary	Free
Public Health	CDC	Public Health in Cities	PLACES	Geographic Portal software services (Geo Portal)	ESRI	Download	Web-based	Proprietary	Free
Environment	EEA / EC	Urban Water management	SIIF - UWWTD	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Open source software	Free
Human Epidemiology	CDC	Human epidemiology: cancer	State Cancer Profiles	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free
Human Epidemiology	CDC	Human epidemiology: cancer	U S Cancer Statistics Data Visualization	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free
Public Health	CDC	Public Health: Life expectancy	U.S. Small-area Life Expectancy Estimates Project (USALEEP) interactive map	Geographic Portal software services (Geo Portal)	-	Download	Web-based	Proprietary	Free

Appendix II Complete list of all agencies and supra-national bodies explored with their link. Table 24: Complete list of all agencies, and supra-national bodies explored. Area, Agency, Field, Name of application, Type of application, tool used to provide geospatial service, distribution, licencing and tool accessibility.

Agency	Division / Project name	Link
CDC	Atlas Plus	https://gis.cdc.gov/grasp/nchhstpatlas/maps.html
CDC	Chronic disease indicators (CDI)	https://nccd.cdc.gov/cdi/rdPage.aspx?rdReport=DPH_CDI.ExploreByTopic&islClass=&islTopic=ALC&islYear=
CDC	Disability and Health Data System (DHDS)	https://dhds.cdc.gov/LP?CategoryId=BARRIER&IndicatorId=CBARRIER&ShowFootnotes=true&View=Map&yearId=YR4&stratCatId1=DISSTAT&stratId1=DISABL&stratCatId2=&stratId2=&responseId=YESNO01&dataValueTypeId=AGEADJPREV&MapClassifierId=quantile&MapClassifierCount=5
CDC	Division of Bacterial Diseases - NCCDPHP	https://www.cdc.gov/gis/gis-at-cdc.htm
CDC	Division of Global Migration and Quarantine	https://www.cdc.gov/gis/gis-at-cdc.htm
CDC	Division of Population Health - NCCDPHP	https://www.cdc.gov/gis/gis-at-cdc.htm
CDC	Division of Reproductive Health - NCCDPHP CVC	https://www.cdc.gov/gis/gis-at-cdc.htm
CDC	Epi Info™	https://www.cdc.gov/epiinfo/index.html
CDC	GRASP	https://www.atsdr.cdc.gov/placeandhealth/about_grasp.html
CDC	Interactive Atlas of Heart Disease and Stroke	https://nccd.cdc.gov/DHDSPAtlas/Default.aspx?state=County&tt=HI&classid=3&subid=8&hifilters=%5B%5B9,1%5D,%5B2,1%5D,%5B3,3%5D,%5B4,1%5D,%5B7,1%5D%5D&ol=%5B10,14%5D
CDC	Interactive Atlas of Heart Disease and Strokes CDC	https://nccd.cdc.gov/DHDSPAtlas/?state=County
CDC	Interactive Diabetes Atlas	https://gis.cdc.gov/grasp/diabetes/DiabetesAtlas.html
CDC	Nutrition, Physical Activity, and Obesity: Data, Trends and Maps	https://nccd.cdc.gov/dnpao_dtm/rdPage.aspx?rdReport=DNPAO_DTM.ExploreByTopic&islClass=OWS&islTopic=OWS1&islYear=20202020
CDC	PLACES	https://experience.arcgis.com/experience/22c7182a162d45788dd52a2362f8ed65
CDC	State Cancer Profiles	https://statecancerprofiles.cancer.gov/map/map.withimage.php?00&county&001&001&00&0&01&0&1&5&0#results

CDC	U.S. Cancer Statistics Data Visualization	https://gis.cdc.gov/Cancer/USCS/?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcancer%2Fdataviz%2Findex.htm#/AtAGlance/
CDC	Nutrition, Physical Activity, and Obesity: Data, Trends and Maps	https://www.cdc.gov/nchs/data-visualization/life-expectancy/
EC	IPCHEM	https://ipchem.jrc.ec.europa.eu/#
EC	Data Europa	https://data.europa.eu/en
EC	INSPIRE Geoportal	https://inspire-geoportal.ec.europa.eu/
ECDC	ECDC Map Maker Tool (EMMa)	https://geoportal.ecdc.europa.eu/mapmaker/?session=1647859481570
ECDC	ECDC Geoportal	https://geoportal.ecdc.europa.eu/geocatalogue/?isUrlSearch=true&session=4ugdhu3zipou
ECDC	ECDC Geocatalogue	https://geoportal.ecdc.europa.eu/geocatalogue/?isUrlSearch=true&session=4ugdhu3zipou
EEA	Disco Map	https://discomap.eea.europa.eu/Index/
EEA	SDI - geospatial data catalogue	https://sdi.eea.europa.eu/
EEA	SIIF - UWWTD	https://uwwtd.eu/
Eurostat	IMAGE	https://ec.europa.eu/eurostat/web/products-eurostat-news/-/wdn-20210218-1
Eurostat	GISCO	https://ec.europa.eu/eurostat/web/gisco
FDA	FDA GeoPortal	https://www.accessdata.fda.gov/scripts/fdatrack/view/track_project.cfm?program=operations&id=operations-OCM-FDA-GeoWeb
JRC	Big Data Analytics Platform (BDAP)	https://jeodpp.jrc.ec.europa.eu/bdap/
NCI	GIS portal for cancer research	https://gis.cancer.gov/canceratlas/app/
USDA	Animal and Plant Health Inspection Service (APHIS)	https://www.aphis.usda.gov/aphis/maps/aphis/aphis-gis
USDA	Global Agricultural & Disaster Assessment System (GADAS)	https://geo.fas.usda.gov/GADAS/index.html

AppendixIII Final search string for retrieving reviews for scoping review

The final search string was constructed as: GIS AND Practices concept AND Environment/Public Health/etc. concept (only reviews + years, language and countries filters)

FINAL SEARCH (WoS CC, BCI, CABI, FSTA, MEDLINE)

Table 25: Search strings

Set	Strings	Notes	Results (2021-09-01)
#12	#7 AND #2 AND #1 **Refined by: DOCUMENT TYPES:** (REVIEW) AND **PUBLICATION YEARS:** (2021 OR 2020 OR 2019 OR 2018 OR 2017) AND **LANGUAGES:** (ENGLISH) AND **COUNTRIES/REGIONS:** (USA OR NORWAY OR FINLAND OR ENGLAND OR BULGARIA OR ITALY OR CANADA OR ROMANIA OR ESTONIA OR GERMANY OR ICELAND OR UK OR LUXEMBOURG OR FRANCE OR RUSSIA OR SPAIN OR NETHERLANDS OR HUNGARY OR SLOVAKIA OR SERBIA OR SWITZERLAND OR UKRAINE OR POLAND OR WALES OR AUSTRIA OR SWEDEN OR CROATIA OR SCOTLAND OR IRELAND OR PORTUGAL OR CZECH REPUBLIC OR LITHUANIA OR DENMARK OR BELGIUM OR NORTH IRELAND OR MONTENEGRO OR GREECE) *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	**+ Country filter**	**530**
#11	#7 AND #2 AND #1 **Refined by: DOCUMENT TYPES:** (REVIEW) AND **PUBLICATION YEARS:** (2021 OR 2020 OR 2019 OR 2018 OR 2017) AND **LANGUAGES:** (ENGLISH) *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	**+ English language only**	880
#10	#7 AND #2 AND #1 **Refined by: DOCUMENT TYPES:** (REVIEW) AND **PUBLICATION YEARS:** (2021 OR 2020 OR 2019 OR 2018 OR 2017) *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	**+ Last 5 years filter**	899
#9	#7 AND #2 AND #1 **Refined by: DOCUMENT TYPES:** (REVIEW)	**+ Only Reviews**	2,353

	Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years *Search language=Auto*		
#8	#7 AND #2 AND #1 *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	**GIS, Practice, Different fields**	112,588
#7	#6 OR #5 OR #4 OR #3 *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	Different fields	42,274,690
#6	TS=(Climat* OR "Global warming" OR Humidity OR Metereolog* OR Precipitation* OR Temperature* OR Weather) *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	Climate	7,417,938
#5	TS=(Environment* OR Agricultur* OR Agroecosystem* OR Air OR Biodiversity OR Biological diversity OR Biomass OR Carbon footprint OR Coast OR Coasts OR Crop OR Crops OR Ecological OR Ecosystem* OR Flood* OR Forest OR Forests OR Habitat* OR Hydrologic* OR Land OR Natural resource* OR Ocean* OR Pest control OR Pollution* OR Sea OR Sustainab* OR Soil OR Species diversity OR Vegetation OR Waste OR Water) *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	Environment	19,490,406
#4	TS=(Contagio* OR Epidem* OR Endem* OR Foodborne OR (Health NEAR/3 (community OR public OR promotion OR system OR "care system" OR surveillance)) OR "Insect vectors" OR Infect* OR Disease* OR Outbreak* OR (Medicine NEAR/3 (occupation* OR preventi*)) OR Pandem* OR Waterborne OR Zoonos* OR Zoonot*) *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	Public Health/ epidemiology	22,218,258
#3	TS=(((Food OR Foods OR feed OR feeds OR foodstuff*) NEAR/3 (safety OR "risk management" OR "Risk assessment*" OR Hygien* OR salubrity OR contaminat* OR microbiology OR microorganism* OR inspect* OR analysis OR "Critical Control Point" OR ccp OR supply OR supplies OR chain OR	Food safety	4,246,638

	chains)) OR Exposure OR Consumption OR Contaminants OR "Emerging risks" OR "Biological hazards") *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*		
#2	TS=("Approach" OR "Approaches" OR "Practice" OR "Practices" OR "Practise" OR "Practises" OR "Method" OR "Methods" OR "Process" OR "System" OR "Systems" OR "Model" OR "Models" OR "Scenario") *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	Practice	55,201,559
#1	TS=("Geographic information system*" OR "Geographical information system*" OR GIS OR ((Geospatial OR "Geo spatial" OR "Spatial data" OR "Spatial information") NEAR/3 (analys* OR application*))) *Databases= WOS, BCI, CABI, FSTA, MEDLINE Timespan=All years* *Search language=Auto*	GIS	152,507

10. References

Agyeman PC, Ahado SK, Boruvka L, Biney JKM, Sarkodie VYO, Kebonye NM and Kingsley J, 2021. Trend analysis of global usage of digital soil mapping models in the prediction of potentially toxic elements in soil/sediments: a bibliometric review. Environmental Geochemistry and Health, 43:1715-1739. doi: 10.1007/s10653-020-00742-9

APHIS, online. Protecting Animal and Plant Health Using Geographic Information Systems.

Bergquist R and Manda S, 2019. The world in your hands: GeoHealth then and now. Geospatial Health, 14:3-16. doi: 10.4081/gh.2019.779

Bloom MS, 2019. Environmental Epidemiology. In: Nriagu J (ed.). Encyclopedia of Environmental Health, Elsevier. pp. 419-427.

Bronzwaer S, Geervliet M, Hugas M and Url B, 2021. EFSA's expertise supports One Health policy needs. EFSA Journal, 19:e190501.

Broomhead T, Ballas D and Baker SR, 2019. Application of geographic information systems and simulation modelling to dental public health: Where next? Community Dentistry and Oral Epidemiology, 47:1-11. doi: 10.1111/cdoe.12437

Card KG, Lachowsky NJ, Althoff KN, Schafer K, Hogg RS and Montaner JSG, 2019. A systematic review of the geospatial barriers to antiretroviral initiation, adherence and viral suppression among people living with HIV. Sexual Health, 16:1-17. doi: 10.1071/sh18104

CDC CfDCaP, online. What is Epidemiology? [Accessed: 1 June 2022].

CDC CfDCaP, online. Geospatial Research, Analysis, and Services Program (GRASP).

Chaix B, 2018. Mobile Sensing in Environmental Health and Neighborhood Research. In: Fielding JE, Brownson RC and Green LW (eds.). Annual Review of Public Health, Vol 39. pp. 367-384.

Cuadros DF, Li J, Musuka G and Awad SF, 2021. Spatial epidemiology of diabetes: Methods and insights. World Journal of Diabetes, 12:1042-1056. doi: 10.4239/wjd.v12.i7.1042

Delgado JA, Short NM, Jr., Roberts DP and Vandenberg B, 2019. Big Data Analysis for Sustainable Agriculture on a Geospatial Cloud Framework. Frontiers in Sustainable Food Systems, 3. doi: 10.3389/fsufs.2019.00054

Dhewantara PW, Lau CL, Allan KJ, Hu W, Zhang W, Mamun AA and Soares Magalhaes RJ, 2019. Spatial epidemiological approaches to inform leptospirosis surveillance and control: A systematic review and critical appraisal of methods. Zoonoses and Public Health, 66:185-206. doi: 10.1111/zph.12549

Dietrich D, Dekova R, Davy S, Fahrni G and Geissbuehler A, 2018. Applications of Space Technologies to Global Health: Scoping Review. Journal of Medical Internet Research, 20. doi: 10.2196/jmir.9458

ECDC ECfDPaC, online. ECDC Map Maker tool (EMMa).

EsriEUSummit, 2021. 2021 GIS in the European Union Summit. Place Esri, Esri.

EUSPA EUAftSP, online. What is GNSS?

Fatima M, O'Keefe KJ, Wei W, Arshad S and Gruebner O, 2021. Geospatial Analysis of COVID-19: A Scoping Review. International Journal of Environmental Research and Public Health, 18. doi: 10.3390/ijerph18052336

Franch-Pardo I, Desjardins MR, Barea-Navarro I and Cerda A, 2021. A review of GIS methodologies to analyze the dynamics of COVID-19 in the second half of 2020. Transactions in Gis, 00: 1–49. doi: 10.1111/tgis.12792

Hollings T, Burgman M, van Andel M, Gilbert M, Robinson T and Robinson A, 2018. How do you find the green sheep? A critical review of the use of remotely sensed imagery to detect and count animals. Methods in Ecology and Evolution, 9:881-892. doi: 10.1111/2041-210x.12973

InspireDirective IOCTFfNS, 2011. Technical Guidance for the implementation of INSPIRE Discovery

Agyeman PC, Ahado SK, Boruvka L, Biney JKM, Sarkodie VYO, Kebonye NM and Kingsley J, 2021. Trend analysis of global usage of digital soil mapping models in the prediction of potentially toxic elements in soil/sediments: a bibliometric review. Environmental Geochemistry and Health, 43:1715-1739. doi: 10.1007/s10653-020-00742-9

APHIS AaPHIS, online. Protecting Animal and Plant Health Using Geographic Information Systems. Available online: https://www.aphis.usda.gov/aphis/maps/aphis/aphis-gis [Accessed: 1 June 2022].

Bergquist R and Manda S, 2019. The world in your hands: GeoHealth then and now. Geospatial Health, 14:3-16. doi: 10.4081/gh.2019.779

Bloom MS, 2019. Environmental Epidemiology. In: Nriagu J (ed.). Encyclopedia of Environmental Health, Elsevier. pp. 419-427.

Bronzwaer S, Geervliet M, Hugas M and Url B, 2021. EFSA's expertise supports One Health policy needs. EFSA Journal, 19:e190501. doi: https://doi.org/10.2903/j.efsa.2021.e190501

Broomhead T, Ballas D and Baker SR, 2019. Application of geographic information systems and simulation modelling to dental public health: Where next? Community Dentistry and Oral Epidemiology, 47:1-11. doi: 10.1111/cdoe.12437

Card KG, Lachowsky NJ, Althoff KN, Schafer K, Hogg RS and Montaner JSG, 2019. A systematic review of the geospatial barriers to antiretroviral initiation, adherence and viral suppression among people living with HIV. Sexual Health, 16:1-17. doi: 10.1071/sh18104

CDC CfDCaP, online. What is Epidemiology? Available online: https://www.cdc.gov/careerpaths/k12teacherroadmap/epidemiology.html -:~:text=By%20definition%2C%20epidemiology%20is%20the,state%2C%20country%2C%20global). [Accessed: 1 June 2022].

CDC CfDCaP, online. Geospatial Research, Analysis, and Services Program (GRASP). Available online: https://www.aphis.usda.gov/aphis/maps/aphis/aphis-gis [Accessed: 3 May 2022].

CDC CfDCaP, online. Interactive Atlas of Heart Disease and Stroke. Available online: https://nccd.cdc.gov/DHDSPAtlas/Default.aspx?state=County&tt=HI&classid=3&subid=8&hifilters=%5B%5B9,1%5D,%5B2,1%5D,%5B3,3%5D,%5B4,1%5D,%5B7,1%5D%5D&ol=%5B10,14%5D [Accessed: 5 February 2022].

Chaix B, 2018. Mobile Sensing in Environmental Health and Neighborhood Research. In: Fielding JE, Brownson RC and Green LW (eds.). Annual Review of Public Health, Vol 39. pp. 367-384.

Cuadros DF, Li J, Musuka G and Awad SF, 2021. Spatial epidemiology of diabetes: Methods and insights. World Journal of Diabetes, 12:1042-1056. doi: 10.4239/wjd.v12.i7.1042

Delgado JA, Short NM, Jr., Roberts DP and Vandenberg B, 2019. Big Data Analysis for Sustainable Agriculture on a Geospatial Cloud Framework. Frontiers in Sustainable Food Systems, 3. doi: 10.3389/fsufs.2019.00054

Dhewantara PW, Lau CL, Allan KJ, Hu W, Zhang W, Mamun AA and Soares Magalhaes RJ, 2019. Spatial epidemiological approaches to inform leptospirosis surveillance and control: A systematic review and critical appraisal of methods. Zoonoses and Public Health, 66:185-206. doi: 10.1111/zph.12549

Dietrich D, Dekova R, Davy S, Fahrni G and Geissbuehler A, 2018. Applications of Space Technologies to Global Health: Scoping Review. Journal of Medical Internet Research, 20. doi: 10.2196/jmir.9458

ECDC ECfDPaC, online. ECDC Map Maker tool (EMMa). Available online: https://www.ecdc.europa.eu/en/publications-data/ecdc-map-maker-tool-emma [Accessed: 9 January 2022].

ESRI, online. GIS Dictionary. Available online: https://support.esri.com/en/other-resources/gis-dictionary/term/49d6945b-f2b1-4039-9cab-29a478c3ab88 -:~:text=A,load%20at%20a%20gauging%20station [Accessed: 1 June 2022].

ESRI, online. What is geocoding? Available online: https://desktop.arcgis.com/en/arcmap/latest/manage-data/geocoding/what-is-geocoding.htm [Accessed: 13 June 2022].

EsriEUSummit, 2021. 2021 GIS in the European Union Summit. Place Esri, Esri.

EUSPA EUAftSP, online. What is GNSS? Available online: https://www.euspa.europa.eu/european-space/eu-space-programme/what-gnss [Accessed: 1 June 2022E].

Fatima M, O'Keefe KJ, Wei W, Arshad S and Gruebner O, 2021. Geospatial Analysis of COVID-19: A Scoping Review. International Journal of Environmental Research and Public Health, 18. doi: 10.3390/ijerph18052336

Franch-Pardo I, Desjardins MR, Barea-Navarro I and Cerda A, 2021. A review of GIS methodologies to analyze the dynamics of COVID-19 in the second half of 2020. Transactions in Gis, 00: 1–49. doi: 10.1111/tgis.12792

Hollings T, Burgman M, van Andel M, Gilbert M, Robinson T and Robinson A, 2018. How do you find the green sheep? A critical review of the use of remotely sensed imagery to detect and count animals. Methods in Ecology and Evolution, 9:881-892. doi: 10.1111/2041-210x.12973

InspireDirective IOCTFfNS, 2011. Technical Guidance for the implementation of INSPIRE Discovery

Services, Version 3.0 Available online: https://inspire.ec.europa.eu/documents/Network_Services/TechnicalGuidance_DiscoveryServices_v3.0.pdf

IPCHEM EC, online. Information Platform for Chemical Monitoring (IPCHEM). Available online: https://ipchem.jrc.ec.europa.eu/ [Accessed: 25, 07, 2022].

Jia P, Cheng X, Xue H and Wang Y, 2017. Applications of geographic information systems (GIS) data and methods in obesity-related research. Obesity Reviews, 18:400-411. doi: 10.1111/obr.12495

Jia P, Stein A, James P, Brownson RC, Wu T, Xiao Q, Wang L, Sabel CE and Wang Y, 2019a. Earth Observation: Investigating Noncommunicable Diseases from Space. In: Fielding JE (ed.). Annual Review of Public Health, Vol 40. pp. 85-104.

Jia P, Xue H, Yin L, Stein A, Wang M and Wang Y, 2019b. Spatial Technologies in Obesity Research: Current Applications and Future Promise. Trends in Endocrinology and Metabolism, 30:211-223. doi: 10.1016/j.tem.2018.12.003

Kanankege KST, Alvarez J, Zhang L and Perez AM, 2020. An Introductory Framework for Choosing Spatiotemporal Analytical Tools in Population-Level Eco-Epidemiological Research. Frontiers in Veterinary Science, 7. doi: 10.3389/fvets.2020.00339

Kirby RS, Delmelle E and Eberth JM, 2017. Advances in spatial epidemiology and geographic information systems. Annals of Epidemiology, 27:1-9. doi: 10.1016/j.annepidem.2016.12.001

Korycinski RW, Tennant BL, Cawley MA, Bloodgood B, Oh AY and Berrigan D, 2018. Geospatial approaches to cancer control and population sciences at the United States cancer centers. Cancer Causes & Control, 29:371-377. doi: 10.1007/s10552-018-1009-0

Kost GJ, 2019. Geospatial Science and Point-of-Care Testing: Creating Solutions for Population Access, Emergencies, Outbreaks, and Disasters. Frontiers in Public Health, 7. doi: 10.3389/fpubh.2019.00329

Lutkevich B (TechTarget), online. software license. Available online: https://www.techtarget.com/searchcio/definition/software-license [Accessed: 3 August].

Malone JB, Bergquist R, Martins M and Luvall JC, 2019. Use of Geospatial Surveillance and Response Systems for Vector-Borne Diseases in the Elimination Phase. Tropical Medicine and Infectious Disease, 4. doi: 10.3390/tropicalmed4010015

Movebank, online. Movebank for animal tracking data. Available online: https://www.movebank.org/cms/movebank-main [Accessed: 8 April 2022].

mvtec, online. IMAGE CLASSIFICATION. Available online: https://www.mvtec.com/technologies/deep-learning/deep-learning-methods/image-classification [Accessed: 9 June 2022].

NASA, online. What Is a Satellite? Available online: https://www.nasa.gov/audience/forstudents/5-8/features/nasa-knows/what-is-a-satellite-58.html [Accessed: 3 August 2022].

Odhiambo JN, Kalinda C, Macharia PM, Snow RW and Sartorius B, 2020. Spatial and spatio-temporal methods for mapping malaria risk: a systematic review. Bmj Global Health, 5. doi: 10.1136/bmjgh-2020-002919

OGC OGC, online. Geocoding API SWG. Available online: https://www.ogc.org/projects/groups/geocodeapiswg [Accessed: 9 June 2022].

OGC OGC, online. Geography Markup Language. Available online: https://www.ogc.org/standards/gml [Accessed: 9 June 2022].

Sahar L, Foster SL, Sherman RL, Henry KA, Goldberg DW, Stinchcomb DG and Bauer JE, 2019. GIScience and cancer: State of the art and trends for cancer surveillance and epidemiology. Cancer, 125:2544-2560. doi: 10.1002/cncr.32052

Samuels-Kalow ME and Camargo CA, Jr., 2019. The Use of Geographic Data to Improve Asthma Care Delivery and Population Health. Clinics in Chest Medicine, 40:209-225. doi: 10.1016/j.ccm.2018.10.012

Seidel DP, Dougherty E, Carlson C and Getz WM, 2018. Ecological metrics and methods for GPS movement data. International Journal of Geographical Information Science, 32:2272-2293. doi: 10.1080/13658816.2018.1498097

Snyk, online. What Is a Software License? Available online: https://snyk.io/learn/what-is-a-software-license/ -:~:text=A%20software%20license%20is%20a%20contract%20between%20the,against%20them%20that%20may%20arise%20from%20its%20use [Accessed: 9 June 2022].

TechTerms, online. Open Source. Available online: https://techterms.com/definition/opensource -:~:text=When%20a%20software%20program%20is,than%20by%20a%20single%20organization [Accessed: 9 June 2022].

TechTerms, online. Software. Available online: https://techterms.com/definition/software [Accessed: 1 June 2022].

TechTerms, online. Hardware. Available online: https://techterms.com/definition/hardware [Accessed: 1 June 2022].

Wang F, 2020. Why public health needs GIS: a methodological overview. Annals of Gis, 26:1-12. doi: 10.1080/19475683.2019.1702099

WHO WHO, online. Public health services. Available online: https://www.euro.who.int/en/health-topics/Health-systems/public-health-services -

:~:text=Public%20Health%20is%20defined%20as,Acheson%2C%201988%3B%20WHO). [Accessed: 1 June 2022].

Wikipedia, online. Proprietary software. Available online: https://en.wikipedia.org/wiki/Proprietary_software [Accessed: 9 June 2022].

Wilkins EL, Morris MA, Radley D and Griffiths C, 2017. Using Geographic Information Systems to measure retail food environments: Discussion of methodological considerations and a proposed reporting checklist (Geo-FERN). Health & Place, 44:110-117. doi: 10.1016/j.healthplace.2017.01.008

11. Abbreviations

2SFCA	2-step Floating Catchment Area
ADIS	European Commission Animal Diseases Information System
AgCROS	Agricultural Collaborative Research Outcome System
AI	Avian Influenza
AMR	Antimicrobial Resistance
APHIS	Animal and Plant Health Inspection Service
API	Application Programming Interfaces
ASF	African Swine Fever
BCI	BIOSIS Citation Index
BIOHAW	Biological Hazards and Animal Health and Welfare
BYM	Besag, York and Mollie Models
CAB	Centre for Agriculture and Bioscience International
CDC	Centers for Disease Control and Prevention
CORINE	Coordinaion of Information on the Environment
COVID-19	Corona Virus Disease
DestineE	Destination Earth Initiative
DG	Directorates-general departments
DWH	Datawarehouse
EC	European Commission
ECDC	European Centre for Disease Prevention and Control
ECMWF	European Centre for Medium-Range Weather Forecasts
EEA	Environmental European Agency
EFSA	European Food Safety Authority
EMMa	ECDC Map Maker Tool
EO	Earth Observation Satellites
ESA	European Space Agency
ESDAC	European Commission Soil Data Centre
EU	Europe
EUBEE	Partnership Prototype Platform on Bee Health
EUROSTAT	European Statistical Office
EUSPA	European Union Agency for the Space Agency
FAO	Food and Agriculture Organization
FDA	Food and Drug Administration
FSF	Free Software Foundation
FSTA	Food Science and Technology Abstracts
GADAS	Global Agricultural & Disaster Assessment System
GBIF	Global Biodiversity Information Facility
GDA	Geospatial Data Act of 2018
GIS	Geospatial Information System
GLM	General Linear Models

GML	Geography Markup Language
GMO	Genetically Modified Organism
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
GRASP	CDC’s Geospatial Research, Analysis, and Services Program
GUI	Graphical User Interface
HIV	Human Immunodeficiency Virus
HYSPLIT	Hybrid Single Particle Lagrangian Integrated Trajectory Model
iDATA	Integrated Data
IPCHEM	Information Platform for Chemical Monitoring
JRC	Joint Research Centre
KDE	kernel density estimation
LGPL	GNU Lesser General Public License
MAUP	Modifiable Aerial Unit Problem
MEDLINE	Medical Literature Analysis and Retrieval System Online
MESE	Methodology and Scientific Support
MS	Member States
NASA	National Aeronautics and Space Administration
NCCDPHP	CDC's National Center for Chronic Disease Prevention and Health Promotion
NCI	U.S. National Cancer Institute
NUTS	Nomenclature of Territorial Units for Statistics
OGC	Open Geospatial Consortium
OIE	World Organisation for Animal Health
OsGEO	Open Source Geospatial Foundation
OSI	Open Systems Interconnection
PLANTS	Pesticides Residues & Plant Health
PSI	Public Sector Information
R4EU	R for Europe
RS	Remote Sensing
SEED	Spatial Explicit Environmental Data for the integrated spatial analysis in risk assessments
SIGMA	EFSA project on Animal Disease Data Model
SIIF	Structured Implementation and Information Framework
UAV	Unmanned Aerial Vehicles
UGCoP	Uncertain Geographic Context Problem
US	United States
USDA	U.S. Department of Agriculture
UWWTD	Waste Water Treatment Directive
WAHIS	The World Animal Health Information System
WHO	World Health Organization
WorldClim	Global Climate Data
WoS CC	Web of Science Core Collection