UNIVERSITY OF OSLO
Department of informatics

An Open Source Approach to Improving GIS Implementations in Developing Countries

Master thesis
60 credits

Jan Henrik Øverland
February 2010

An Open Source Approach to Improving GIS Implementations in Developing Countries

Master thesis

Jan Henrik Øverland
University of Oslo
Department of informatics
janhov@ifi.uio.no

February 2010

Abstract

Geographic information system (GIS) implementations have a tendency to fail in developing countries. This thesis seeks to explore whether a web based free/open source software (FOSS) development approach can improve chances for sustainable and successful implementations.

To be able to conclude in this matter I have spent 15 months exploring GIS implementations in developing countries such as Sierra Leone and India, as well as reviewing existing GIS software in the market. In this thesis I identify reasons for why most GIS implementations in developing countries fail, what limitations are present and how FOSS can deal with the technical, economical and practical aspects of them. I also show that FOSS has become sufficiently mature and capable to build rich GIS applications in general. On the technical level I document a personally developed solution that solves a major issue regarding map data for most developing countries. Additionally, I show how technical restrictions to web based GIS can be worked around.

During this process I have utilized FOSS frameworks and tools to develop a GIS application for developing countries. This application is integrated into the District Health Information Software (DHIS), a flexible open source health information system that has gained a strong foothold in developing countries over the last years. In order to succeed I have utilized the technical skills and experience I have acquired through professional training by the founders of the software combined with knowledge from the literature and the developing country field study presented in this thesis. A demo is available at demo.dhis2.org (username *admin* and password *district*).

Acknowledgements

I would like to thank my teaching supervisor Jørn Braa for the feedback during the writing process and for giving me the opportunity to travel abroad to conduct the necessary field research and technical training for this thesis.

Additionally, I want to thank the DHIS 2 core developer team for constituting a great environment for learning and collaboration.

Table of Contents

Figures

1 Introduction

This chapter aims to get the reader up to speed on the background, motivation and research objectives for this thesis.

1.1 Context of this thesis

This thesis is part of an ongoing action research project called Health Information Systems Program (HISP), which aims to develop sustainable computer based Health Information Systems (HIS) to support district management and enable local analysis and action. HISP, over the years, has developed an application called District Health Information System (DHIS), which supports data collection and analysis at all levels of the health administration. This application has been experiencing continuous prototyping and customization to meet the growing needs of health information systems in different countries (Lewis 2005).

HISP started the development of the DHIS in South Africa in 1996 (Braa et al 2002). HISP has since then taken part in the process to reconstruct the health services in South Africa and has developed a district-based health information system including software, standardisation of routine health data and general approaches which is now implemented all over the country. DHIS is still being maintained on an ongoing basis. The South African development has focused on action research, user participation and local involvement in the development of an information system aiming at strengthening local management and decentralisation in the health sector. Empowerment of local health managers through the use of local information for decision making is among the key objectives. The relative success of the health information system in South Africa has led to an export of the DHIS and the approach to health management to several countries in Africa and Asia (Øverland 2006).

Since 1999 DHIS has been customized, adapted and tested in various countries outside of South Africa. Mozambique and India were the first nodes of HISP as they were the first two countries where the initiative of transfer of technology started in the year 1998 and 2000 respectively. The replication and implementation of such a system from one country to another presents various complex problems, not only in technological aspects but also related to social, cultural, political and contextual issues (Lewis 2005).

Because of the foothold Microsoft (MS) had and still has in the government sectors around the world, the initial versions of the DHIS software was developed for Microsoft Windows, using the Access database. As HISP tries to expand the use of DHIS to more users and other countries, new demands for the software arise that are hard or impossible to address with the current software. Limitations of the Access database are starting to show, as some of the deployed systems have larger datasets than the database is intended for. There are demands for integrating DHIS with other health information systems, like electronic patient record systems (EPR), health programs, and competing systems. It must also be easier to integrate DHIS with other applications, like web based analysis tools and GIS software. Many developing countries would like to try out or switch to open, non-proprietary technologies (Nordal 2006).

The DHIS 1.3 implementation effort in Cuba in 2003 revealed certain limitations of the existing application. Technical constraints related to the data model led to skepticism among the users. Even if the system is free and regarded as open source software, it is dependent on proprietary Microsoft technologies. This implies unacceptable expenditure for developing countries when considering the number of licenses needed in order to cover every unit in a national HIS. The system has poor networking support, which limits the ability to scale up across units. Also, the system is programmed in Visual Basic 6; a language of which Microsoft has ended free support and announced will be phased out.

The decision to start developing a new version, the DHIS 2, was taken during the spring of 2004 at the University of Oslo, and emerged out of these circumstances. The system is entirely based on free and open source frameworks, implying no acquisition cost for the health departments adopting it. The system is platform independent and uses an object relational persistence system, which implies that the system will run on any operating systems and on most database management systems. (Øverland 2006).

The author has been involved in HISP and a member of the DHIS 2 core developer team since the summer of 2008. The need for an integrated geographic information system (GIS) has become gradually more important and longed for, and as HISP has seen several GIS projects for the DHIS fail to meet the needed requirements and usability over the years, he was handed the responsibility of creating an integrated web based GIS module that could be distributed within the DHIS2.

1.2 Motivation

My motivation for this project and this thesis is two-fold. Firstly, an encouraging factor is that GIS seem to improve health information systems (HIS). Taylor (1991) claims that GIS can be applied in different fields and in different subjects, including the health sector, where it has a great potential to support decision making. Burrough (1986) states that GIS is a powerful set of tools for storing, retrieving, transforming and displaying spatial data from the real world for a particular set of objectives. Chetley et al (2006) points out that inefficient allocation of scarce resources and lack of coordination among key stakeholders has made duplication of efforts, overlapping responsibilities, and resource wastage common and troublesome problems that may be handled with the presence of GIS. Sauerborn and Karam (2000) illustrate that most of the data contained in health information systems are spatial data, in the sense that the data are tied to a specific area, such as the catchments area of a health centre, a health district or to a geographic point (village or hospital). They say that such data are "geo-referenced" and argue that when data from health information system are fed in to a GIS, it can enhance the HIS in four important areas:

- *Data Communication*
  When data are displayed through "maps" it is easily comprehensible to decision makers, health practitioners, laypersons and the media alike.
- *Data Analysis or spatial analysis*
  GIS can not only display the results of statistical analysis in map format but also can calculate the population in a radius around a water point. This type of calculation or buffering of population can only be done using geographic information system.
- *Decision support*
  The intuitive grasp of maps and the ability to display both health information

system indicators and the results of data analyses makes GIS a valuable tool for decision support.

- *Links to other sectors*

  As maps are a generic platform for displaying information from all sectors, like education, economic development, infrastructure, finances, agriculture etc, they provide an opportunity to share data from different sectors and foster exchange and communication.

These statements fit well into the context of developing countries, where inefficient allocation of scarce resources and lack of coordination, as well as the need for improved primary health care services are, well-documented phenomena.

Secondly, as a software developer I have a personal interest in learning new technologies and expanding my knowledge base. The new web based GIS frameworks that are currently evolving look promising and mastering them might be advantageous in the future.

1.3 Research objectives

This section contains my research objective and questions, which are situated in the context of web based free and open source GIS and GIS implementation in developing countries.

Research objective:
Explore the capability of web based open source GIS development frameworks and open source methodology to improve sustainability of GIS implementations in developing countries.

To answer the research objective I have dispersed the research objective into two questions:

Research question 1:
Explore the technical capability of web based FOSS GIS development frameworks.

Research question 2:

Explore limitations and solutions regarding development and implementation of GIS in developing countries in general, and whether web based FOSS GIS development frameworks and open source methodology can improve sustainability of such implementations.

The experiences and findings from the research questions are assembled and practiced in my own simultaneous GIS project. The evaluation of this project strengthens my conclusion to the research objective of this thesis.

2 Literature review

In this chapter I present the relevant theoretical background for this thesis. Theories and background information will be reflected in the empirical study and discussed in relation to the empirical findings.

2.1 Open source methodology

Open source describes practices in production and development that promote access to the end product's source materials, typically their source code. Some consider open source a philosophy, others consider it a pragmatic methodology. Before the term open source became widely adopted, developers and producers used a variety of phrases to describe the concept; open source gained hold with the rise of a public, worldwide, computer-network system called the Internet, and the attendant need for massive retooling of the computing source code. Opening the source code enabled a self-enhancing diversity of production models, communication paths, and interactive communities. Subsequently, a new, three-word phrase "open source software" was born to describe the environment that the new copyright, licensing, domain, and consumer issues created.

The open source model includes the concept of concurrent yet different agendas and differing approaches in production, in contrast with more centralized models of development such as those typically used in commercial software companies. A main principle and practice of open source software development is peer production by bartering and collaboration, with the end-product (and source-material) available at no cost to the public (Wikipedia M 2010).

Free software is software that can be used, studied, and modified without restriction, and which can be copied and redistributed in modified or unmodified form either without restriction, or with minimal restrictions only to ensure that further recipients can also do these things and that manufacturers of consumer-facing hardware allow user modifications to their hardware. Free software is generally available without charge.

In practice, for software to be distributed as free software, the human-readable form of the program (the source code) must be made available to the recipient along with a notice granting the above permissions. Such a notice either is a "free software license", or a notice that the source code is released into the public domain. The free software movement was conceived in 1983 by Richard Stallman to satisfy the need for and to give the benefit of "software freedom" to computer users. From 1998 onward, alternative terms for free software came into use, with “free and open source software” (“FOSS”) as the most common. The antonym of free software is "proprietary software" or "non-free software". Commercial software may be either free software or proprietary software, contrary to a popular misconception that "commercial software" is a synonym for "proprietary software". An example of commercial free software is Red Hat Linux. Free software, which may or may not be distributed free of charge, is distinct from "freeware" which, by definition, does not require payment for use. The authors or copyright holders of freeware may retain all rights to the software; it is not necessarily permissible to reverse engineer, modify, or redistribute freeware.

Since free software may be freely redistributed it is generally available at little or no cost. Free software business models are usually based on adding value such as applications, support, training, customization, integration, or certification. At the same time, some business models which work with proprietary software are not compatible with free software, such as those that depend on a user paying for a license in order to lawfully use a software product (Wikipedia N 2010).

The most well-known written material on open source methodology is “The Cathedral and the Bazaar” (Raymond 2000) by Eric Raymond. His essay contrasts two different software development models. In the C*athedral* model development takes place in a centralized way. Roles dedicated to design, implementation, and project management are clearly defined and the code are restricted to an exclusive group of software developers. Brooks (1995) advocates this model and says that in order to preserve the architectural integrity of a system, the system design should be done by as few architects as possible. The B*azaar* model, however, is different. Here, roles are not clearly defined and the code is developed over the Internet in view of the public. It is formulated as a series of prescriptions – technically we could call them patterns or principles – that should be applied in order to make a project successful. The principles that are to be used in a project depend on several factors, such as project size (software size and number of developers) and status (whether a project is in its early stages or an already consolidated project with a large user-developer base etc). The application of a specific principle is not very well defined either; its use should be guided by common sense, i.e. adjusted to the circumstances of the project in question (Robles 2004). Robles singles out some of the most vital principles:

- *Treat your users as co-developers*
  This is to open the development process to the maximum, so that developers interested in the project can be integrated seamlessly. More co-developers equals the possibility for the project of evolving more quickly, of creating usable functionality at a faster pace, etc. This principle can be regarded as *participatory design* when it comes to research, which is presented in the next chapter. Schuler (2008) states that the practice of open source communities can be recognized as participatory design (ref 3.2.2).
- *Early releases*
  The first versions of the software have to be released, although limited in functionality, as soon as there is something presentable and functional. This way, the possibility of finding co-developers interested in participating increases (Robles 2004). According to Raymond (2000), these early and frequent releases are a critical part of the open source software development model. Most developers used to believe this was bad policy for larger than trivial projects, because early versions are almost by definition buggy versions and you don't want to wear out the patience of your users. However, given a large enough co-developer base, almost every problem will be characterized quickly.
- *Modularize to the maximum*
  The general structure of the software should be modular allowing for parallel development and reuse of code on independent components.

2.2 ICT, OSS and GIS

Abbreviations: Information and Communication Technology (ICT), Open Source Software (OSS), Geographic Information System (GIS). The debate in the 1990s over choosing between ICT and other development imperatives has now shifted from one of tradeoffs to one of complementary (The World Bank Group 2003). In this section I will present relevant literature from the field of ICT in developing countries.

2.2.1 Status on ICT in developing countries

All United Nations member states have pledged to meet the UN Millennium Development Goals (MDG) by the year 2015. The UN (2000) summarizes the MDG: "The MDG bind countries to do more and join forces in the fight against poverty, illiteracy, hunger, lack of education, gender inequality, child and maternal mortality, disease and environmental degradation. The eight goal (...) calls on rich countries to relieve debt, increase aid and give poor countries fair access to their markets and their technology."

The United Nations Development Programme (UNDP) recognizes that ICT will play a key role in the fight against global poverty and as an effective tool in helping to achieve the MDG. ICT opens for participation in the global markets; it promotes political accountability, improves the deliveries of basic services and enhances local development opportunities (UNDP 2005).

There is however several challenges related to the process of applying ICT for development. The Swiss Agency for Development and Cooperation (SDC) and the Global Knowledge Partnership (GKP 2003) presents some of these challenges: "ICT need to be affordable for the poor, in terms of both initial outlay and on-going costs. Information received needs to be relevant, contextualised and available in the local language. Communication needs to be timely, so that information is obtained or provided neither too soon nor too late. And people require the capacities to use information."

Furthermore SDC and GKP suggest that there is a call for action on three broad areas:

- There is a need to integrate ICT systematically into poverty reduction strategies.
- One needs to move beyond small pilot projects to a larger nation-wide or even region wide implementation of ICT programmes.
- One has to continue to create new types of partnerships involving all major stakeholders – government, civil society and the private sector.

Related to these points is the issue of sustainability; the need for development efforts to last. The UN (2000)Millennium Project points out: "Development is largely an expression of local initiative and international partnership; it cannot be sustained without local ownership and champions."

2.2.2 ICT in organizations

The survival and growth of organizations in increasingly turbulent contemporary environments depends upon effective utilization of ICT for aligning the organizational structure with environmental changes. How ICTs can help organizations in responding to the challenges of effectively harnessing ICTs, to achieve flexible organizational structures are key ongoing challenges for developing countries (Sahay & Avgerou 2002). ICT can play a substantial role in the following major areas:

- Improving access to services
- Strengthening the basis for decision-making
- Promoting information exchange among users
- Enhancing the effectiveness of institutions.

Quality in health care delivery is largely dependent on the availability of and access to information, which directly contributes to the capacity building of the service providers, and increases the awareness and thereby the health seeking behavior of the community. Thus, the use of ICT can help reduce disparities between the services available in urban and rural areas and reduce the costs involved in transporting patients to urban facilities. Because an effective information dissemination system enhances the participation among the stakeholders more than an application, ICT is fundamental to enhancing knowledge, and communicating for better health. ICT itself does not do anything useful; in order to realize any gains, it must become part of an information system. As Heeks (1999) has argued, it is important to emphasize that these technologies only provide new mechanisms for handling an already existing resource: information. Therefore, to understand ICTs, one must first understand information practices and needs.

The increasing global interdependencies and the accelerating pace of change demand more flexible and adaptive organizations (Malone & Crowston 1991). Kenaroglu (2000) has defined organizational flexibility in terms of "vulnerability" and "adaptability". Then, effective implementation of ICT can potentially decrease vulnerability by reducing the cost of expected failures and enhance adaptability by reducing the cost of adjustment. Piore and Sabel (1984) cited in Kenaroglu (2000), argue that ICT-based systems offer organizations the opportunity of functional integration, multi-skilled staff, rapid and flexible decision-making structures. This implies a greater delegation of responsibilities and greater autonomy to operating units, and a more flexible and "organic" approach enabling a quick adjustment to changing environmental conditions (Kenaroglu 2000). Information management skills rely on the ability to make choices about the optimal arrangements for particular situations. Unlike earlier generations of technology, ICT offers not a single "best" way of organizing, but rather represents a set of more or less appropriate alternative organizing, staffing, and managing options that may be adopted in different organizational contexts.

Thus, following Heeks, ICT cannot be understood without analyzing information. Developing an enterprise system requires information about several different things. For example, this means information relating to *supply*, such as the availability and sources of finance, labour, technology, raw materials, and other enterprise inputs. Information is also required about *demand*, including market opportunities and its characteristics such as issues related to location, price, size, and quality. Information is also needed about *other environmental factors*, like competitors, laws, etc (Heeks 1999).

2.2.3 ICT in developing countries

Many researchers see ICT as a powerful new opportunity for at least some developing countries to improve their competitive position in certain fields and to foster their development precisely because of their relative lack of established infrastructure. However, often, the focus is placed entirely on the technology, and not enough on the information and the practices surrounding it that are required to make their ICTs deliver effective outputs.

For instance, in Mozambique, banks, public and private institutions and the government are currently engaged in introducing ICT in order to improve their services. Examples have been shown by Mosse and Sahay (2003), related to ICT in the health domain where they argue that "Mozambique has been attempting to introduce ICTs in various sectors to promote socio-economic development". This is also supported by Macome (2003), when she talks about ICT projects in rural communities, the "Telecentre Project", which is the first experience of its kind in Mozambique.

Although the socio-economic structure of many developing countries are not flexible on handling organizational or institutional changes, the complex interrelations between these changes and information technology have significant implications for the way ICT does and will affect the societies and economies of these countries. The main issue facing developing countries is thus not so much the access to a particular technology, but dealing with the challenges related to the processes of technological change and the human and social factors that need to be adapted to these processes. Also, the introduction of ICTs requires certain new skills of design, maintenance, and management, as well as complementary infrastructural facilities such as reliable telephone systems, power supplies, and physical infrastructure like roads and transport. Deficiencies in these factors prevent the widespread adoption of information technology in developing countries. Quality of data, too, requires an adequate level of skill, infrastructure, and managerial know-how that is generally lacking in developing countries.

These constraints on ICT development in developing countries have been well documented by researchers. For example, Mosse and Sahay (2003), in relation to the introduction of ICT in the health sector, argue that these ICT initiatives take place in a context that is historically and culturally shaped; the socio-cultural structures are reflected in patterns of how work is currently done. NORAD (2002) recognizes the challenges posed by ICT in development contexts and acknowledges the increasing digital gap between the rich and the poor parts of the world. In May 2000 a working group submitted a report entitled "Bridging the digital divide – challenges and opportunities for NORAD and its development partners". In the light of the report NORAD decided to integrate ICT into development cooperation in order to combat poverty more effectively.

2.2.4 OSS in developing countries

"Consequently one major argument against the implementation of proprietary software in the public sector is the subsequent dependency on proprietary software vendors. Whenever the proprietary standards are established the necessity to follow them is given. Even in an open tender acquisition system, this requirement for compatibility with proprietary standards makes the system biased towards specific software vendors, perpetuating a dependency." (Gosh et al 2002).

Software development has traditionally been done following a model with formal division of labour that uses proprietary knowledge, guarded by restrictive intellectual property rights, enclosed within a corporate hierarchy, to guide and govern the process. (Weber 2003)
Software plays an increasingly important role in the global economic markets as well as in most international organisations and non-governmental organisations. The ability of a country, and the firms within it, to interact with these markets and organisations is to a large extent restricted by its ICT capacity. Weber (2003) states that fairly sophisticated information technology should be thought of now as prerequisites to effectively interact with the world economy. This implies that decisions governments take about procurement, standard setting and adoption, technology investments and training is critical.

Due to the digital divide, and more specifically due to the fact that developing countries have limited budgets earmarked for information technology (Weber 2003), most governments in the developing world are advocating the use of FOSS when it is a feasible alternative to proprietary software solutions. Weber lists three identifiable motivations for why developing countries have chosen to embrace the use of FOSS:

- *Independence*
  Many developing countries have acknowledged that they are increasingly dependent on software suppliers located in other countries. The costs associated with implementation, licenses, and maintenance of this proprietary software is high, in addition these services do not nurture the national economy.
  By advocating FOSS, local contractors can compete by price and quality on the delivery of support and maintenance, generating jobs and boosting the economy. Expensive licenses is no longer an issue, in addition the maintenance can be replicable without incurring large costs as the modification of source code is also free (Weber 2003).
- *Security and autonomy*
  One of the proclaimed advantages of FOSS compared to proprietary software is that of security. The main argument is that bugs are generally fewer and when a bug is identified it is fixed much faster in FOSS. In addition, FOSS assures that the software is secure as code can be inspected; governments need to rely on systems without elements controlled by third parties, possibly located outside

the country, posing a threat to national security. Another benefit by introducing FOSS in critical governmental systems is that one will achieve diversity in the technical base which decreases the potential damages caused by computer attacks targeting monolithic code. One of the responsibilities of most governments is to provide free access to public information. The use of standards and open formats instead of data tied to single providers guarantees this free access. Furthermore, to ensure permanence of this data, it is important to be free from the goodwill of single suppliers or monopoly conditions. Developing countries also feel that their influence on how proprietary software developed is very limited, FOSS promises more flexibility and allows autonomous input on software development.

- *Intellectual property rights and productivity*
  The intensified combat against software piracy has led many countries to advocate FOSS as an alternative to expensive proprietary software. Lower costs is one thing, another matter is that of ownership. By choosing FOSS tools, the possibility of utilizing and expanding the software is no longer limited by proprietary rights, the potential of the software tool is now only limited by the knowledge, learning and innovative energy of the users. Extensive use of FOSS will form a technological infrastructure dependent on the delivery of other products and services, potentially boosting local economy. By combining inexpensive technical manpower with free software, local companies in emerging economies can get competitive advantages in both local and global markets.

2.2.5 IS in developing countries

Do most information systems (IS) projects in developing countries succeed or fail? Any attempt to answer this question must start by categorizing success and failure (Heeks 2002).

Defining and measuring success and failure

Any success/failure categorization runs into some immediate difficulties that Heeks' article cannot completely resolve. The first difficulty is the subjectivity of evaluation – viewed from different perspectives, one person's failure may be another's success. The categorization does try to address this within the limits imposed by the subjectivity of the case study writers themselves. The second difficulty is the timing of evaluation - today's IS success may be tomorrow's IS failure, and vice versa (Heeks 2002).

First, there was the total failure of an initiative never implemented or in which a new system was implemented but immediately abandoned. A second possible outcome is the partial failure of an initiative, in which major goals are unattained or in which there are significant undesirable outcomes. In some cases, where only a subset of initially stated objectives has been achieved, the notion of partial failure may be relatively straightforward. Finally, one may see the success of an initiative, in which most stakeholder groups attain their major goals and do not experience significant undesirable outcomes (Heeks 2002).

The extend of success and failure

What proportion of developing country IS projects fall into each of the three outcome categories? No one knows for certain. The question is hard enough to answer in the industrialized countries. There, at least, a certain level of surveys, evaluations, and analysis is present (Korac-Boisvert & Kouzmin 1995, James 1997, Sauer 1999, The Economist 2000). On the basis of the range of figures provided in these surveys [listed in Heeks article], one may estimate that something like one-fifth to one-quarter of industrialized-country IS projects fall into the total failure category; something like one-third to three-fifths fall into the partial failure category; and only a minority fall into the success category. This, at least, can be used as a threshold indicator to answer the question.

What is the evidence relating to IS success and failure in developing countries? Evidence to address the earlier question, and move beyond the threshold estimations just offered, is very limited. In addition to poor recognition of subjectivity and timing of evaluation, the constraints on evidence are several:

- *Lack of literature in general*
  Until very recently, the entire literature on IS and developing countries would struggle to fill a single bookshelf. The attention of writers – from researchers to consultants to journalists – has been focused elsewhere.
- *Lack of evaluation*
  Those who have the will to evaluate – such as academics – often lack the resources and capacity. Those who have the resources – such as aid donor agencies – often lack the will to evaluate.

- *Focus on case studies*
  The literature on IS in developing countries has grown, but it is a literature dominated by case studies of individual IS projects. Taken alone, these provide no basis for estimation of overall failure/success rates.

Despite these limitations, there are some glimpses of evidence. An overview of the literature concludes, "successful examples of computerisation can be found . . . but frustrating stories of systems which failed to fulfill their initial promise are more frequent". (Avgerou & Walsham 2000). A few more specific multiple-case studies have been conducted, with examples summarized here:

- Health information systems in South Africa: Braa and Hedberg (2002) reported widespread partial failure of high cost systems with little use of data.
- IS in the Thai public sector: Kitiyadisai (2000) reported "failure cases seem to be the norm in Thailand at all governmental levels."
- Donor-funded IT projects in China: Baark and Heeks (1999) reported that all were found to be partial failures.
- World Bank-funded IT projects in Africa: Moussa and Schware (1992) reported almost all as partial failures.

Design-actuality gaps

We need to simultaneously evaluate the current system and the future system. Yet, by definition, they cannot simultaneously exist. It is relatively easy to assess the current "actuality" in a location. But in order to assess the future, we must assess instead the representation of an intended future – an intended future that is represented in a design for the system. The model to be used here is therefore based on an assessment of the match or mismatch between local actuality ("where we are now") and system design ("where the design wants to get us"). Put simply, we refer to this as the design–actuality gap (Heeks 2002).

The most extreme form occurs when industrialized country designers create an information system within and for an industrialized-country context, and that IS is subsequently transferred to a developing country. In such situations, the actuality of local conditions in the developing country will not have been considered at all in the original design, and a considerable design–actuality gap is therefore likely, leading to a significant risk of IS failure. Even if some effort is made to develop an information system specifically for a developing country organization, similar problems can arise. Industrialized-country stakeholders, such as consultants or IT vendors or aid donors, often dominate the IS design process in developing countries. Those stakeholders bring their context with them and inscribe it into their IS designs: inscriptions that will mismatch developing country actuality. Some stakeholders bring with them the "If it works for us, it'll work for you" mentality that makes no attempt to differentiate between industrialize d and developing contexts. Others will differentiate, but - given their poor understanding of local developing country conditions - their assumptions about user actuality will be incorrect. In all cases, large design–actuality gaps and high failure risks are the outcome (Heeks 2002).

In practice, because of subjective expectations about the future and subjective perceptions of reality, it could be argued that every individual IS stakeholder has their own design and their own version of actuality. Among these myriad design–actuality gaps, we must necessarily simplify the model. Drawing on another thread within the failure literature (Lyytinen & Hirschheim 1987, Sauer 1999), the two key homogenized stakeholders will be the designers who create the dominant IS design, and the users who populate the local actuality. These groups are especially valuable to an understanding of failure given their dislocation, in both psychological and even physical terms, as part of the IS implementation process. However, this simplification does impose limits – for example, limiting subjective partial failures to a consideration of the objectives of these two stakeholder groups alone (Heeks 2002).

What could be relevant dimensions of this design-actuality gap between the designers' dominant design and the local actuality of the users? The dimensions could be built up in a number of ways: theoretically on the basis of information systems literature; descriptively on the basis of a straightforward delineation of components of an information system; and analytically on the basis of case studies.

Furthermore, Heeks states that the design is a representation of an intentional future. It is a world-in-miniature that contains elements that have been inscribed either explicitly or implicitly. These elements include:

- *Components from the designers' own context*
  IS design is a situated action—an action "taken in the context of particular, concrete circumstances" (Suchman 1987). This action draws elements of that context into the design: "Our technologies mirror our societies. They reproduce and embody the complex interplay of professional, technical, economic and political factors" (Bijker & Law 1992). Designers themselves are part of and shaped by that context, so their own cultural values, objectives, etc. will be found inscribed in the design (Shields & Servaes 1989, Braa & Hedberg 2002).
- *Conceived assumptions about the situation of the user*
  This includes assumptions about the users' activities, skills, culture, and objectives, and assumptions about the user organization's structure, infrastructure, etc. (Boehm 1981, Suchman 1987, Clemons et al 1995, Wynn & deLyra 2000).

He explains that the most extreme form occurs when industrialized country designers create an information system within and for an industrialized-country context, and that IS is subsequently transferred to a developing country. In such situations, the actuality of local conditions in the developing country will not have been considered at all in the original design, and a considerable design–actuality gap is therefore likely, leading to a significant risk of IS failure.

Even if some effort is made to develop an information system specifically for a developing country organization, similar problems can arise. Industrialized-country stakeholders, such as consultants or IT vendors or aid donors, often dominate the IS design process in developing countries. Those stakeholders bring their context with them and, even if located in a developing country, they will inscribe that context into their IS designs: inscriptions that will mismatch developing country actuality. Some stakeholders bring with them the "If it works for us, it'll work for you" mentality that makes no attempt to differentiate between industrialized and developing contexts. Others will differentiate, but - given their poor understanding of local developing country conditions - their assumptions about user actuality will be incorrect. In all cases, large design–actuality gaps and high failure risks are the outcome.

Combined with more descriptive material on information systems, these theoretical ideas build to create seven dimensions of relevance to design–actuality gaps:

information (data stores, data flows, etc.); technology (both hardware and software); processes (the activities of users and others); objectives and values (the key dimension, through which factors such as culture and politics are manifest); staffing and skills (both the quantitative and qualitative aspects of competencies); management systems and structures; and other resources (particularly time and money) (Heeks 2002).

The contexts of designer and user are often distant in physical, cultural, economic, and many other ways. The remoteness of designers means that their contextual inscriptions are liable to be significantly different from user actuality. So, too, are the inscribed assumptions that remote designers make about that actuality. Design–actuality gaps are therefore more extreme and more explicit and, as a result, are easier to identify and to understand. Developing country cases therefore provide valuable data that helps illuminate both IS failure and underlying structures and processes. Put another way, developing country cases make it easier to move beyond the black box (Akrich 1992).

These theories, design-actuality gaps in particular, are essential to the evaluation of my own project. This is discussed in eighth chapter.

2.2.6 GIS in developing countries

The term GIS describes computerized information storage, processing and retrieval systems that are specifically designed to cope with geographically-referenced spatial data and the corresponding attribute information. These systems have the potential to support activities of organizations in managing spatially distributed resources by examining trends, identifying factors that cause them, revealing alternative paths to solve a problem, and indicating the implication of decisions.

Al-Romaithi (1997) describes GIS technology as a product of the developed world, which has unique complexities and problems when applied to developing countries because of their very different socio-economic realities and priorities. Cultural differences in concepts of time, scale, detail, distance, values, topology and relationships mean that GIS implementation is context sensitive. Beyond these cultural differences, GIS implementation is also affected by institutional contexts and organizational interrelationships (Martin 1998).

Thus, the implementation of GIS in non-western settings requires a flexible and context sensitive approach (Martin 1998), involving a variety of modifications to suit local needs. Successful investigations of GIS installations in non-western contexts require an approach that analyses the interactions between the technology and the specific social or institutional setting.

Despite these multiple difficulties in effectively applying GIS, its value in developing countries is becoming increasingly significant, given the current worldwide concern about the state of our environment and the pressure to sustainably manage natural resources. For example, in recent years GIS technology has been used in health care setting of developing countries to:

- Organize and analyze information. There is a growing understanding and appreciation regarding the power of health and health-related information in planning and implementing health programs. Health information is also becoming more and more readily available. Given these points and the fact that most health information is tied in some way to geography, it is becoming increasingly important that health professionals, organizations, and communities create systems that empower them to really take advantage of the many different types of information that is available and that can be brought to bear on health issues and program management.
- Assist in planning and implementing, but it is also a powerful tool to present ideas and motivate people to take action: GIS and maps in general can be a powerful tool when presenting ideas as many people learn best with visual aids. Remember, one picture can be worth a thousand words. Presenting ideas using maps and GIS can help people see patterns and to better understand service gaps or barriers to access. For example, many regions in the world have very rugged terrain or limited transportation routes. Without systems that allow you to present information on service sites and information about geographic barriers simultaneously, it is possible that you could misinterpret the information available.
- Design more carefully target health programs to specific population needs: It is crucial that health programs are tailored to the specific needs and unique characteristics of a community. GIS allows you to characterize and organize information about a community and link those characteristics to the services that

are provided. What languages are spoken in a community? What is the distribution of those living in the community by age, and gender? What communities have the highest number of women of childbearing age? Where are mosquitoes most prevalent? Is there clean drinking water? What percentage of the population is in poverty? Systems that allow you to organize and analyze this data geographically can be very powerful when designing health programs and assessing health needs.

- Track and monitor the incidence of disease and/or inventorying available health resources: Understanding and monitoring the incidence of disease or conducting inventories of existing resources is a simple and powerful use of GIS. Where is a disease most prevalent? What population has the greatest disease burden? Where existing health resources are located and where are gaps in services? These are classic uses of GIS and can be easily implemented.

2.2.7 GIS and health care

GIS and related spatial analysis methods provide a set of tools for describing and understanding the changing spatial organization of health care, for examining its relationship to health outcomes and access, and for exploring how the delivery of health care can be improved. Although GIS has been used for several decades to examine health care systems the scope of GIS contributions has grown rapidly in recent years. Advances in computing power and graphics, as well as the development of GIS based location analysis models and methods have stimulated innovative health care applications (McLafferty 2003).

GIS has been increasingly used to map and explore geographical variation in need for health services and to develop innovative indicators of healthcare need (McLafferty 2003). Due to its (GIS) advantage of spatial database management and display capabilities, it has been used to link diverse layers of population and environmental information to characterize the many dimensions of healthcare need for small areas (Mohan 1993). One such example is the effort at creating “community environmental health profiles” that describe demographic, economic, and lifestyle characteristics of the population and also exposures to potential environmental hazards (Peters and Hall 1999). Although the layering capabilities of GIS have been used for many years, researchers are now making use of the analytic capabilities to relate data sets that rely on non-consistent area units and to generate meaningful service areas (Mohan 1993, Lovett et al 1998).

Access to health care is an important issue in most of the countries. Access describes people's ability to use health services when and where they are needed (Aday and Anderson 1981). GIS helps to put emphasis on the geographical dimensions of access. Healthcare decisions are strongly influenced by the type and quality of services available in the local area and the distance, time and cost of traveling to reach those services. GIS has been used to create better measures of geographical access and to analyze geographical inequalities in access as well as those patterned along social and economic lines (McLafferty 2003).

2.3 ICT and GIS implementation for public health in developing countries

It is now accepted that the use of computers in the health systems of developing countries is "a need, not a fashion" (Sepulveda et al 1992). Studies have shown that microcomputers not only lead to an improvement in the quality of decision-making and to more efficient and rational management of resources, but that they also bring about a significant reduction in the costs of data-processing (Sandiford et al 1992).

Nowadays there is an evolution regarding the diffusion of ICT, and consequently GIS, within many developing countries. Most of these disseminations aim at speeding up research and development processes through regulatory reforms, thus accelerating community access to new promising products, providing tools for better decision-making support, evaluations and benchmarking, addressing inequities, and enhancing monitoring capability for governments.

The information revolution, of which GIS is an integral part, is taking place in society, and embedded in a broader context of socio-economic change. The socio-economic realities and priorities of the "third world" are quite different and, if GIS is to be used for the challenges facing developing countries, then it must respond to those realities and priorities. However, we must carefully evaluate why and how a technology should be adopted before "jumping onto the technology bandwagon." How technology is applied in a health care environment may have serious legal consequences for those involved. Factors such as privacy of information, consent, liability, jurisdiction, and other, are issues that come to mind while introducing ICT and GIS technologies within health care in developing countries. One important ongoing area of application of ICTs in developing countries concerns the public health domain (Saugene 2005).

2.3.1 How can ICTs support health information systems in developing countries?

Wilson et al (2001) define HIS as a set of tools and procedures that a health programme uses to collect, process, transmit, and use data for monitoring, evaluating and controlling the health system. Health management is a pre-requisite for effective health services, and can potentially be improved by better HIS. However, HIS in developing countries face diverse problems such as resource (human and infrastructure) constraints, poor information, multiplicity of programs, and donor dependence. Lippeveld and Sapirie (2000) argue that most developing countries have routine paper-based health information systems in place to collect and report data. These are seen largely inadequate and ineffective to support health care.

Braa et al (2004) argues that HIS in developing countries tend to be data-led where data is seen an end in itself, rather than being action-led, with a focus on how collected information will inform decision and action.

ICT has the potential to change the delivery of health care services and patient care, and the management of the health care system around the world. Technologies and applications are changing at ever increasing speeds and so are the dynamics of the process surrounding the implementation of e-health technologies and applications. Some important areas of applications concern accelerating patient access to new and promising technology. As a result, most developing countries are attempting to strengthen and computerize their health information systems, but most of them have to date yielded unsatisfactory results. This is because the implementation of HIS in developing countries is a complex and very challenging task, as the process demands not only a technology transfer, but also the introduction of a different kind of culture that accompanies the system. In addition, public health setting in developing countries is a complex environment, characterized by the existence of different donors, different levels of organization, and use of top-down approaches for decision-making.

Basically, there are two things to help make HIS work effectively in challenging developing countries contexts. The first one concerns sustainability, which refers to how the HIS can work in practice, over time and in a local setting. This involves shaping and adapting the systems to a given context, cultivating local learning processes and institutionalizing routines of use that persist over time. The other challenge refers to scalability which concerns the problem of how to make one working solution spread to other sites, and be successfully adapted there (Braa et al 2004). To support the sustainability and scalability of health information systems; it is of vital importance to generate local, self-sufficient learning processes together with working mechanisms for the distribution of appropriately formatted experiences across sites. It is also emphasized that interventions must be aligned with the surrounding configurations of existing institutions, competing projects and efforts, as well as with every day practices. In relation to the introduction of ICT in health sector in Mozambique, Mosse and Sahay (2003) advocate that historical and cultural practices and socio-cultural structures are reflected in patterns of how work is currently done.

2.3.2 Challenges when introducing GIS systems

Even though the introduction of GIS is very similar to other ICTs, it would appear that GIS is particularly difficult to handle from an organizational viewpoint. Budic and David (1994) addresses this by arguing that introducing computerized technology, such as GIS into local government operations is risky business. They support the arguments by underlining that implementation obstacles are reported even by successful users of GIS technology. However, a couple of factors have to be taken into account while developing a GIS system. Al-Romaithi (1997) adds that for GIS to be of use to addressing the current challenges facing developing countries, planners would need to actively respond to these realities and priorities, including such problems as: inadequate financial and human resources, infrastructure bottlenecks, difficulties in initiating change in government organizations, the existing work-culture, and issues related to the transfer of appropriate technology. Sahay and Walsham (1997) points out that one of the main reasons for failure in GIS implementations is that very little attention is given to adapting the technology to the needs and capabilities of the countries and organizations in which it is going to be adapted.

Many different factors have been described to contribute to project failure, which we have synthesized under four broad categories (Sahay & Walsham 1997):

- Technology transfer
- Institutional factors
- Data management
- Manpower

Technology transfer

GIS technology is a product of the developed world and is generally introduced into the context of developing countries through the process of "technology transfer." In majority of cases, the transfer of GIS technology to developing countries is facilitated by international aid agencies. As a result, the process of technology transfer is influenced significantly by the agenda and management styles of these specific agencies involved.

Differences of opinion between the aid-agencies sponsoring GIS projects and officials of the recipient country are often responsible for the breakdown of the transfer process. Aid projects normally come with stipulations about the kind of software and methodologies that should be used, and consultants from developed countries are called in to oversee project management.

Problems associated with technology transfer are related to the contents of what is transferred, the structure of the transfer process, and the absence of mechanisms to sustain the system once the aid project is completed. For a truly sustainable transference of GIS technology, it is important to ensure that within the domain of the receivers of technology, which are often state institutions, conditions are established wherein project can be continued and reinforced (Sahay & Walsham 1997). This process of technology transfer is again fraught with a number of significant problems related to institutions, manpower, data, and project management.

Institutional factors

Like any new technology or tool GIS can be intoxicating. Difficulties with this can be technological but many times may be organizational. Various authors have highlighted a number of institutional factors, especially within government organizations, that significantly influence the effective use of GIS in developing countries. Sahay and

Walsham (1997) point to the problems that arise because of the existing culture within government agencies. If users are not actively involved in the process of change, the project can be a disaster. Often people are determined to implement complicated GIS software and present fancy maps, but do not realize how time and resource intensive they can be to create and maintain overtime. Choosing the right software and learning how to use it, ensuring that you have the appropriate hardware to run the software, finding and entering good data, ensuring that all your data is compatible and in the right format are issues that can individually or collectively complicate and stall GIS projects even for those who are very experienced. To make decisions using GIS, there has to be good cooperation between the computer specialist who is developing the system, and the subject-expert who has to interpret the output. However, the developer and the user are typically responsible to different ministries and departments, and the functional manner in which these organizations operate makes the sharing of data and other technical and organizational resources extremely problematic (Sahay & Walsham 1997).

Sometimes it is easier and more efficient just to use a map with a set of pins or to place dots on a simple computer generated map. Sometimes the benefits of creating a GIS do not out-weigh the costs. Like with all projects, it is important to clearly understand what you hope to accomplish and have a good sense of the questions that you want to answer before you jump in. Once you know this, you can accumulate the data and decide what tools you need to power your system. However, sometimes it can be problematic for an organization to handle the long data establishment period. This often leads to frustration because of the long period between the investment and the realization of the promised value.

Data management factors

Like all data systems, GIS is only as good as the data that powers the system or the expertise or experience of the person implementing the system. One picture can be worth a thousand words but it does not mean that you are painting the correct picture. For example, if you create a set of maps that shows that all of a regions health facilities are clustered in one area you could be convinced that the area has ample resources. However, if you do not know the population of that area, the expertise of the providers in the clinics, and/or the actual burden of disease in the area then you cannot draw any clear conclusions regarding the appropriateness of the resource distribution. As a result, the effectiveness of GIS depends on the degree of relevant data as inputs. One of the most challenging problems with GIS is finding timely data in the appropriate format. An effective implementation of GIS is severely vulnerable due to the limited availability of useful geographical data. Two aspects lead to this problem, namely: the existence of data and the accessibility of existing data.

The first problem is concerned with maps which are scarce because making maps as well as updating them is a costly and time consuming activity and in many developing countries is difficult due to financial constraints. In addition to information from maps, information about physical and socioeconomic features also tends to be scarce, because obtaining these types of data need field surveys which are time consuming and expensive.

An issue that also contributes to the non-availability of appropriate spatial data arises from cultural and technical limitations. This is also influenced by the dependence of many GIS applications in developing countries on data generated using remote-sensing technology. The aspect regarding the lack of accessibility of data is derived from the fact that different existing data sets tend to be hard to combine. In some cases creating a combination of data sets even proves to be impossible (Longley et al 2001, Teeffelen et al 1993). One reason for this is that often the newly established GIS projects function as isolated islands of innovation. Many software packages have built-in mechanisms that allow you to link databases in other formats or import data from other programs. Sometimes this is easy but often it is difficult. You can always enter your own data into the system but this can be extremely time-consuming, particularly when you are mapping multiple layers of demographic information about a regions or community. Technical data problems also arise because of data being collected in non-standardized formats, whose conversion is not supported by standard GIS software. Absence of policies to define data standards for access and exchange magnifies the problem of developing GIS systems, and many GIS projects are initiated without any coherent data management strategies.

3 Method

In this chapter I give an account of the research methods and research approach that have been conducted.

3.1 Research methods

This section describes the research methods.

3.1.1 Participatory action research

Essentially participatory action research is research which involves all relevant parties in actively examining together current action in order to change and improve it. They do this by critically reflecting on the historical, political, cultural, economic, geographic and other contexts which make sense of it. Participatory action research is not just research which is hoped that will be followed by action. It is action which is researched, changed and re-researched, within the research process by participants. Nor is it simply an exotic variant of consultation. Instead, it aims to be active co-research, by and for those to be helped. Nor can it be used by one group of people to get another group of people to do what is thought best for them - whether that is to implement a central policy or an organisational or service change. Instead it tries to be a genuinely democratic or non-coercive process whereby those to be helped, determine the purposes and outcomes of their own inquiry (Wadsworth 1998).

According to Baskerville (1999), participation is fundamental to action research; it is an approach which demands that participants perceive the need to change and are willing to play an active part in the research and the change process. All research requires willing subjects, but the level of commitment required in an action research study goes beyond simply agreeing to answer questions or be observed. Here, the responsibility for theorizing is shared with client participants, members of the organization who are actively engaged in the quest for information and ideas to guide their future actions (Baskerville 1999). During the research and development process of my project end users participated and contributed to design and functionality for the prototypes and to my search for limitations and pitfalls regarding GIS implementation in developing countries.

Greenwood and Levin (1998) say that the action research process is built upon three key elements; research, action, and participation. As in action research, the researchers and the organizational people/stakeholders work together to solve a problem. The method seeks not primarily to look for generalizations, but concentrates on solving real life problems while creating new practical or theoretical knowledge. In traditional social research where interaction with what you are studying is not emphasized, action research strongly focuses on this interaction. The researchers must participate with the stakeholders in the specific context to obtain insights that cannot be understood when studying it from a distance. And it is assumed that the researcher cannot acquire the depth of understanding that the problem owner will have already achieved through years of living within the social context under study (Baskerville 1999). Action researchers believe that everyone has the potential to analyze their own situation, contribute to the process and add valuable knowledge and understanding to the others involved. The core of this principle is democracy, since the inclusion of the local stakeholder as co-researchers democratizes the research process. The different aspects of the context and problem are well known by the local stakeholders, and together with the professional researcher this can be addressed with appropriate methods. The researcher will gain knowledge from the other participants about the problem, and the health staff will learn about ways to deal with these problems with the researcher. Action researchers bring their knowledge of action research and general information system theories, and the local stakeholder brings situated, practical knowledge into the action research process. This way, participatory action research is based on assumptions that reality is situated and social systems are self-referencing (Baskerville 1999). In the GIS project chapter we will see that not only end users, as mentioned earlier, but also HISP members and people at the WHO have been involved in the process.

Greenwood and Levin (1998) defines action research as social research carried out by a team encompassing a professional action researcher and members of an organisation or community seeking to improve their situation. Action research promotes broad participation in the research process and supports action leading to a more just or satisfying situation for the stakeholders. Thus, action research is thus a process of social research where both outsiders and problem owners work together to solve problems and reach common goals. The method seeks not primarily to look for generalizations or to provide theories to be true or false, but concentrates on solving real life problems while creating knowledge. Action research focuses on the interaction between the researcher and the stakeholders, contrary to conventional or critical research. Participation with the problem owners in the specific context is seen as necessary to obtain insight in matters that cannot be understood when studying it from a distance. The professional researcher can also be seen as an actor who is able to loosen up tensions between the stakeholders or break up a situation they may be stuck in (Greenwood and Levin 1998). Some groups of people can have communication problems and the role as a neutral outsider may help the researcher to act as a broker and to address these problems.

3.1.2 Action research in the field of IS

Action research has been an established research method in the social sciences since the mid-twentieth century, but it was not until toward the end of the 1990s it began to become popular for use in scholarly investigations of information systems (Baskerville 1999). Following The International Federation for Information Processing conference in 1998, Avison et al (1999) reported that AR and other qualitative approaches have now gained acceptance at the same level as quantitative studies in the field of IS. In 2003, the shift to qualitative methods by mainstream researchers was further manifested with a special issue of the prestigious paper MIS Quarterly named "Action Research in Information Systems".

Braa et al (2004) define a perspective on action research which they call networks of action. They argue that in efforts to institutionalize and make sustainable changes, actions needs to be situated in networks rather than on singular units. The establishment of such networks of actions creates opportunities for sharing experience, knowledge, technology, and value between the various nodes of the experience. Scalability is thus a prerequisite for sustainability of local action (Braa et al 2004). The amount of knowledge transfer that has taken place is important – when the researcher leave, the local teams are responsible for continuing the work. Networks of these teams that can struggle together and learn from each other provide more sustainable projects than singular units (Nordal 2006).

3.1.3 Case study

According to Cornford and Smithson (1996) a case study is an in-depth exploration of one situation. This exploration often needs to have a certain time span, as a snapshot of a situation at a particular moment can not capture the process of change. In a case study the researchers devote themselves to the specific situation, and the reward is a richness of data, obtained by multiple means. A single case study can be hard to use for generalisations, but by finding other similar studies this can be addressed by developing stronger relationships for certain relationships.

Walsham (1993) denotes case studies as "interpretive", where the various researchers may have different perceptions of a study. The purpose of the different studies is to reveal "a truth" rather than "the truth", since a case study will be interpreted differently amongst people, and the presentation of the case will be based on the researchers perception of the phenomena described.

A frequent criticism of the case study methodology is that its dependence on a single case makes it incapable of providing a generalizing conclusion. Giddens (1984) considered case methodology "microscopic" because it lacked a sufficient number of cases. He argued that the relative size of the sample independent of the number of cases being used does not transform a multiple case into a macroscopic study. The goal of the study should establish the parameters, and then should be applied to all research. In this way, even a single case could be considered acceptable, provided it met the established objective (Øverland 2006).

3.2 Research approach

In this section I firstly describe the HISP team, as my research for this thesis has taken place exclusively through being part of it. Secondly, I describe the part of the open source methodology (ref 2.1) that is related to research; the participatory design approach.

3.2.1 The HISP team

HISP has requirements for systems in many areas. The organisational structure needs to reflect these requirements. It aims at optimizing utilization of the technical competence in the network by organizing the members in international groups. The structure has two dimensions: The first one is based on geography, referred to as national teams. The second is based on technical focus, referred to as technical teams.

National teams (with number of members):

- Global team (11, mainly Oslo based)
- Indian national team (13)
- Vietnamese national team (9)
- West African international team (3)
- Tanzanian national team (3)
- Malawian national team (4)

Technical teams:

- Core team
- GIS team (where I am the leader)
- Patient/community team
- Mobile team
- Integration team
- Infrastructure team

The DHIS2 leadership and coordination efforts are centralized in Oslo, but local development teams have been established in the countries HISP are piloting DHIS2. These have their own coordinators that supervised by Oslo. My role has so far included global development of GIS. The plan is to expand the GIS team in the future, which means that I will be responsible for training and team coordination as well. The history of HISP is described in the introduction chapter.

3.2.2 User participation: Participatory design in HIS development

Participatory design is an approach to design that attempts to actively involve all stakeholders (e.g. employees, partners, customers, citizens, end users) in the design process to help ensure that the product designed meets their needs and is usable (Wikipedia K 2010).

There are many approaches to system design and implementation. Some of these involve complex methodologies that have been developed for large commercial systems development projects. Other approaches are based on constructing working models or "prototypes" of the system to be worked with and evaluated by the end-user. Still others involve combinations of these approaches with varying degrees of emphasis on study, documentation and prototypes. However, selecting the best approach depends on many factors including: level of support, previous experience with automated information systems, existing information technology policy and practices, potential costs of implementation, and so on. Approaches to implementation can vary. They can be large and comprehensive or begin with relatively simple and inexpensive applications - initial emphasis is placed on learning - and become more complex and operations oriented as time progresses. This section will highlight approaches that focus on the role of users in the development process (Saugene 2005).

In the traditional systems analysis methodology, the importance of user involvement was frequently stressed. However, the computer professional was the person who was making the real decisions and driving the development process. Systems analysts were trained in, and knowledgeable of, the technological and economic aspects of computer applications but far more rarely on the human (or behavioural) aspects which are at least as important. The end-user, the persons who are going to use the system, frequently felt resentment, and top management did little more than pay lip service to computing. The systems analyst may be happy with the system when it is implemented. It may conform to what the system analysts understand are the requirements, and does so efficiently. However, this is of little significance if the users, who are the customers, are not satisfied with it. The participatory approach is aimed at changing users' perspectives about the information system. They would see the information system as a tool for the skilled worker, and the worker should be in control of the tool (Saugene 2005).

Problems within the health information system are heavily laden with cultural, political, and economical values. Greenbaum and Kyng (1991) argue that when information systems are introduced within an organization they change the organization. The design of these changes needs to start with an understanding of the use situation. But, the HIS network is not easily describable; we cannot expect the information system analyst, who sees the workplace from the outside, to capture the same vision about the organization as someone involved in the day-to-day activities.

The social perspective on IS has been popular among many Scandinavian IS researchers, and several Scandinavian IS projects have contributed to this research area. User participation in system development is the main component of the approach. This refers to the involvement of users in different activities in the system development process.

Following RHINO (2001), the restructuring of routine health information systems should involve all key stakeholders in the design process. Experience suggests that systems that are designed by a team of "information experts" without adequate involvement of key stakeholders usually fail to reflect the needs and practical reality of service providers and managers, and do not encourage ownership of the system.

Many actors have pointed out reasons for user participation such as:

- To improve the knowledge upon which systems are built
- To enable people to develop realistic expectations and reduce resistance to change
- To increase workplace democracies by giving the members of an organization the rights to participate in decisions that are likely to affect their work.

The first two reasons are targeted at using knowledge of the workers to tailor the new system to the actual work it is meant to support. The third reason however is more related to cultural and political aspects of systems development, aiming at improved workplace democracy. This participatory approach will change the user perspective about the information systems. They would see the information systems as a tool for the skilled worker, and the worker should be in control of the tool. Involving users as competent practitioners in the HIS change process will help avoid wide range of visions about the new technology (Saugene 2005).

The starting point of the Scandinavian design (participatory) approach is that every human should have the right to participate equally in decisions concerning his or her life. This is about the importance of inclusion of skilled users in the process of design and use of computer-based information systems (Bjørn-Andersen & Hedberg 1997, Ehn 1993). This approach is politically significant, interdisciplinary, and action-oriented. It raises questions on democracy, power, and control at the workplace and assumes that the participation of skilled users in the design process can contribute importantly to successful design and a high quality product (Ehn 1993).

Many projects were developed with the aim of finding strategies for increasing work life democracy through user participation in the system development effort, i.e. involving users in work activities during systems development by giving them the rights to participate in decision making in areas that are likely to affect their work.

In order to bring the same view of the workplace and to ensure sustainability of health information systems, Greenbaum and Kyng (1991) argues that "involving the users on design process can help on selection and better understanding of the problem and this will lead to mutual learning between this and the IS analyst about their respective fields by providing same perspectives of work." This will also help in creating a common vision about the new technology among users and between the users and the developers.

However, the restructuring of routine health information systems should involve all key stakeholders in the design process. Experience suggests that systems that are designed by a team of "information experts" without adequate involvement of key stakeholders usually fail to reflect the needs and practical reality of service providers and managers, and does not encourage ownership of the system. Assessment and design of routine health information approaches should involve a broad range of stakeholders, including representatives of all management levels of the health system (RHINO 2001).

Interviews

Interviews offer researchers the chance to explore topics in depth and to gain appreciation of the context within which the interview was conducted (Cornford & Smithson, 1996). Some of the interviews I conducted in Sierra Leone turned out to be very useful as they revealed a lot of limitations and gave me valuable insight in order to understand the design context. There are essentially three types of interviews (Examiner 2009):

- *Structured interviews*

  Structured interviews require adherence to a very particular set of rules. Each question that is outlined should be read word for word by the researcher without any deviation from the protocol. In some cases, the interviewer is also required to show consistency in behavior across all interviews. This includes bodily posture, facial expressions, and emotional affect. Reactions to participant responses should be kept to a minimum or avoided entirely. Structured interviews are the type used most often by quantitative researchers. The style is most useful when looking for very specific information. The benefits are that it keeps the data concise and reduces researcher bias.

- *Semi-structured interviews*
  Semi-structured interviews are a bit more relaxed than structured interviews. While researchers using this type are still expected to cover every question in the protocol, they have some wiggle room to explore participant responses by asking for clarification or additional information. Interviewers also have the freedom to be more friendly and sociable. Semi-structured interviews are most often used in qualitative studies. The style is most useful when one is investigating a topic that is very personal to participants. Benefits include the ability to gain rapport and participants' trust, as well as a deeper understanding of responses. Data sets obtained using this style will larger than those with structured interviews.

- *Unstructured interviews*
  Unstructured interviews have the most relaxed rules of the three. In this type, researchers need only a checklist of topics to be covered during the interview. There is no order and no script. The interaction between the participant and the researcher is more like a conversation than an interview. Unstructured interviews are most often used in ethnographies and case studies (types of qualitative studies). They are best used when researchers want to find as much information as possible about their topic. The benefit is that unstructured interviews often uncover information that would not have been exposed using structured or semi-structured interviews. The researcher and participant are not limited by the protocol. Data sets collected using unstructured interviews will be larger than the rest (Examiner 2009).

Personally, I found unstructured interview the most satisfactory. As mentioned, a benefit of unstructured interviews is that they often uncover information that would not have been exposed using structured or semi-structured interviews, and I wanted the interview objects to open up and fill me in on HIS limitations in developing countries, expectations for the new GIS etc.

Questionnaire

During the workshop in India I handed out a questionnaire that hopefully would bring in valuable suggestions and information. People were supposed to draw on their knowledge from both the health sector and the IT sector in developing countries and present their thoughts on what they expect from a GIS for the DHIS. This information collection approach turned out to be, in this case, rather useless as the majority considered themselves unable to contribute because they had no experience with GIS at all. They felt that they did not know what they were talking about and thus that their answers could simply be misleading information.

Development and collaboration

By developing a web based GIS application myself I have gained a lot of skills and knowledge on the relevant development frameworks that I definitely would not have by simply looking into the topic. Collaborating with specialists and professionals obviously motivated and engaged me even more. This is explained in the GIS project chapter.

Prototyping

The prototype method was formally introduced to the information system community in the early 1980s to combat the weakness of the traditional waterfall model. The early prototyping process was for developers to design and build a scaled-down functional model of the desired system and then the developers demonstrate the working model to the user. This results in comments and feedback on its suitability and effectiveness. The developer then continues to develop the prototype until the developers and the users agree that the prototype is satisfactory (Avison & Fitzgerald 2003).

Prototyping is an important part of Rapid Application Development (RAD) and is used to help establish the user requirements and in some cases the prototype evolves to become the system itself. Prototyping helps to speed up the process of eliciting requirements, and speed is obviously important in RAD, but it also fits the RAD view of evolving requirements and users not knowing exactly what they want until they see or experience the system (Avison & Fitzgerald 2003).

Prototyping addresses some of the problems of traditional systems analysis; in particular, the complaint that users only see their information system at implementation time, when it is too late to make changes. With prototyping, user acceptance of a system is regarded as far more likely. By implementing a prototype first, the analyst can show the users something tangible – inputs, intermediary stages, and outputs – before finally committing to the new design.

A prototype is frequently built using special tools such as screen painters and report generators, which facilitate the quick design of screens and reports. The user may be able to quickly see what outputs will look like. Prototyping is also regarded as a way of encouraging user participation. The hands-on use of prototyping by users provides experience, understanding, and the opportunity for evaluation. Once users and managers realize that things could be changed and that they could exert influence, it can lead to improved participation and commitment to the project (Saugene 2005).

4 The Sierra Leonean context

This chapter provides a background of the Sierra Leonean context, which is necessary in order to understand the project incidents and important when analyzing the research objectives.

4.1 Health

All medical care is generally charged for in Sierra Leone (Kambia 2010) and is provided by a mixture of government, private and non-governmental organizations (NGOs). There are over 100 NGOs operating in the health care sector in Sierra Leone. The Ministry of Health and Sanitation is responsible for organising health care and after the end of the civil war the ministry changed to a decentralised structure of health provision to try and increase its coverage.

The country is divided into 13 health districts that correspond to the districts of Sierra Leone except for the Western Area Rural and Western Area Urban districts which are combined into the Western Area Health district. Each district has a health management team and an average of 50 peripheral health units (PHU) and over 100 technical staff. The management team is responsible for planning, organising and monitoring health provision, training personnel, working with communities and supplying equipment and drugs. As a part of the recovery after the civil war, new health clinics open every week throughout the country.

The PHUs are designed to be the delivery point for primary health care in the country and there are three main types. The *community health centre* carries out health prevention measures, cures and health promotion activities and is in charge of overseeing the other PHUs in the area. It is planned that all chiefdoms, the unit of local government in Sierra Leone below the level of district, should have at least one community health centre. *Community health post*s perform a similar function to community health centres, but have fewer facilities and are used to refer patients to the health centre or the district hospital. *Maternal and Child Health post*s are the first level of contact on the ground and are located in smaller towns of with populations between 500-2000. Much of the health care infrastructure was decimated during the Civil War and the health service is still in the process of being organised with hospitals and PHU being rebuilt or created and staff being trained (Wikipedia B 2010). Public health in Sierra Leone is generally poor and in 2007 the country had the highest level of child mortality in the world (Walsh 2008).

4.2 Politics

Sierra Leone is a constitutional republic with a directly elected president and a unicameral legislature. The current system of government in Sierra Leone, established under the 1991 Constitution, is modeled on the following structure of government: the Legislature, the Executive and the Judiciary. Within the confines of the 1991 Constitution, supreme legislative powers are vested in Parliament, which is the law making body of the nation. Supreme executive authority rests in the president and members of his cabinet and judicial power with the judiciary of which the Chief Justice is head. The president is the head of state, the head of government and the commander-in-chief of the Sierra Leone Armed Forces and the Sierra Leone Police. The president appoints and heads a cabinet of ministers, which must be approved by the Parliament. The president is elected by popular vote to a maximum of two five-year terms. The president is the highest and most influential position within the government of Sierra Leone.

The current parliament in the August 2007 Parliamentary elections is made up of three political parties with the following representations; the All People's Congress (APC) 59 seats, the Sierra Leone People's Party (SLPP) 43 seats, and the Peoples Movement for Democratic Change (PMDC) 10 seats. The most recent parliamentary elections were held on August 11, 2007. The All People's Congress (APC), won 59 of 112 parliamentary seats; the Sierra Leone People's Party (SLPP) won 43; and the People's Movement for Democratic Change (PMDC) won 10. To be qualified as Member of Parliament, the person must be a citizen of Sierra Leone, must be at least 21 years old, must be able to speak, read and write the English language with a degree of proficiency to enable him to actively take part in proceedings in Parliament; and must not have any criminal conviction.

Since independence in 1961, Sierra Leone's politics has been dominated by two major political parties, the Sierra Leone People's Party (SLPP), and the ruling All People's

Congress (APC), although other minor political parties have also existed but with no significant supports.

The judicial power of Sierra Leone is vested in the judiciary, headed by the Chief Justice and comprising the Sierra Leone Supreme Court, which is the highest court in the country and its ruling therefore cannot be appealed; High Court of Justice; the Court of Appeal; the magistrate courts; and traditional courts in rural villages. The president appoints and parliament approves Justices for the three courts. The Judiciary have jurisdiction in all civil and criminal matters throughout the country (Wikipedia A 2010). The Sierra Leone Ministry of Foreign Affairs and International Relations, headed by Minister of Foreign Affairs Zainab Hawa Bangura is responsible for foreign policy of Sierra Leone. Sierra Leone has diplomatic relations that include China, Libya, Iran, and Cuba. Sierra Leone has good relations with the West, including the United States and has maintained historical ties with the United Kingdom and other former British colonies through membership of the Commonwealth of Nations.

Former President Siaka Stevens' government had sought closer relations with other West African countries under the Economic Community of West African States (ECOWAS) a policy continued by the current. Sierra Leone, along with Liberia and Guinea form the Mano River Union (MRU) primarily designed to implement development projects and promote regional economic integration between the three countries. Sierra Leone is a member of the United Nations and its specialized agencies, the African Union and the African Development Bank (AFDB) as well as the International Criminal Court with a Bilateral Immunity Agreement of protection for the US military (Wikipedia A 2010).

4.3 Economy

Sierra Leone is slowly emerging from a protracted civil war and is showing signs of a successful transition. Investor and consumer confidence continue to rise, adding impetus to the country's economic recovery. There is greater freedom of movement and the successful re-habitation and resettlement of residential areas.

Rich in minerals, Sierra Leone has relied on mining, especially diamonds, for its economic base. The country is among the top 10 diamond producing nations in the world. Mineral exports remain the main foreign currency earner. Sierra Leone is a major producer of gem-quality diamonds. Though rich in diamonds, it has historically struggled to manage their exploitation and export.

Annual production of Sierra Leone's diamond estimates range between $250–300 million U.S dollar. Some of that is smuggled, where it is possibly used for money laundering or financing illicit activities. Formal exports have dramatically improved since the civil war with efforts to improve the management of them having some success. In October 2000, a UN-approved certification system for exporting diamonds from the country was put in place and led to a dramatic increase in legal exports. In 2001, the government created a mining community development fund, which returns a portion of diamond export taxes to diamond mining communities. The fund was created to raise local communities' stake in the legal diamond trade.

Sierra Leone is perhaps best known for its blood diamonds that were mined and sold for high prices during the civil war. In the 1970s and early 1980s, economic growth rate slowed because of a decline in the mining sector and increasing corruption among government officials.

By the 1990s economic activity was declining and economic infrastructure had become seriously degraded. Over the next decade much of the formal economy was destroyed in the country's civil war. Since the end of hostilities in January 2002, massive infusions of outside assistance have helped Sierra Leone begin to recover. Much of the recovery will depend on the success of the government's efforts to limit corruption by officials, which many feel was the chief cause for the civil war. A key indicator of success will be the effectiveness of government management of its diamond sector. By 2008, Sierra Leone had a gross domestic product (GDP) per capita of around $900 (CIA 2010).

Sierra Leone has one of the world's largest deposits of rutile, a titanium ore used as paint pigment and welding rod coatings. Sierra Rutile Limited, owned by a consortium of United States and European investors, began commercial mining operations near the city of Bonthe, in the Southern Province, in early 1979. It was then the largest non-petroleum US investment in West Africa. In 1990, the company and the government made a new agreement on the terms of the company's concession in Sierra Leone. Rutile

and bauxite mining operations were suspended when rebels invaded the mining sites in 1995, but exports resumed in 2005.
About two-thirds of the population engages in subsistence agriculture, which accounts for 52.5% of national income. The government is trying to increase food and cash crop production and upgrade small farmer skills. The government works with several foreign donors to operate integrated rural development and agricultural projects.
Despite its successes and development, the Sierra Leone economy still faces significant challenges. There is high unemployment, particularly among the youth and ex-combatants. Authorities have been slow to implement reforms in the civil service, and the pace of the privatisation programme is also slacking and donors have urged its advancement (Wikipedia A 2010).

4.4 Demographics

The 2009 UN estimate of Sierra Leone's population is 5 696 000. Freetown, with an estimated population of 1 070 200, is the capital, largest city and the hub of the economy, commercial, educational and cultural centre of the country. Bo is the second city with an estimated population of 269 000. Other cities with a population over 100 000 are Kenema, Koidu Town and Makeni.

Although English is the official language spoken at schools, government administration and by the media, Krio (language derived from English and several African languages and native to the Sierra Leone Krio people) is the most widely spoken language in virtually all parts of Sierra Leone. The Krio language is spoken by 98% of the country's population and unites all the different ethnic groups, especially in their trade and interaction with each other.
According to the World Refugee Survey 2008, published by the U.S. Committee for Refugees and Immigrants, Sierra Leone had a population of 8 700 refugees and asylum seekers at the end of 2007. Nearly 20 000 Liberian refugees voluntarily returned to Liberia over the course of 2007. Of the refugees remaining in Sierra Leone, nearly all were Liberian (World Refugee Survey 2009).

The life expectancy of Sierra Leone is 41 years (Mackenzie 2007).

4.5 Education

Education in Sierra Leone is legally required for all children for six years at primary level and three years in junior secondary education (Wang 2007), but a shortage of schools and teachers has made implementation impossible (U.S. Department of Labour 2002). Two thirds of the adult population of the country are illiterate (UNDP 2009). The Sierra Leone Civil War resulted in the destruction of 1,270 primary schools and in 2001 67 percent of all school-age children were out of school (U.S. Department of Labour 2002). The situation has improved considerably since then with primary school enrollment doubling between 2001 and 2005 and the reconstruction of many schools since the end of the war (Wang 2007). Students at primary schools are usually 6 to 12 years old, and in secondary schools 13 to 18. Primary education is free and compulsory in government-sponsored public schools.

The country has three universities, the University of Sierra Leone, founded as Fourah Bay College in 1827 (the oldest university in West Africa), and Njala University, primarily located in Bo District, which was established as the Njala Agricultural Experimental Station in 1910 and became a university in 2005 (SL Encyclopedia 2008). Teacher training colleges and religious seminaries are found in many parts of the country.

4.6 ICT policies

This section describes national and educational ICT policies.

4.6.1 National

A national policy on ICT is almost non-existent at the present time. However the policymaking process began in 2006 and it is expected to be finalised in 2007. A Telecommunications Act of 2006 has, however, been passed and has set the pace for the establishment of a regulator the National Telecommunications Commission (NaTCom), with responsibility for licensing and spectrum management among other things.

4.6.2 Educational

The absence of a national ICT policy has equally affected the ICT in education policy. However, provisions for ICT utilization are embedded in the National Science and

Technology Policy, with assertions such as making science and technology education compulsory in the basic education system by integrating it into the curricula of all schools and at all levels. The policy also states that “the rapid development and exploitation of ICTs shall be targeted.”

At the same time, the National Education Master Plan 1997–2006 outlines plans for upgrading teachers through the use of distance education. In support of distance education and learning aided by ICTs, the government’s reform initiatives include restructuring and upgrading of the School Broadcasting Unit in support of the 6-3-3-4 system of education. An upgraded Educational Broadcasting Division has been proposed to replace the School Broadcasting Unit. One of the objectives of the new proposed division is to produce and deliver quality educational radio and television programmes to complement and enrich lessons in formal and non-formal education classes.
The following table provides a framework for understanding the core factors that help and hinder the development of ICTs in education in Sierra Leone (Mangesi 2007).

Factors	**Enabling Features**	**Constraints**
Policy framework and implementation plans	A commitment in the Education Master Plan	Lack of a national and educational ICT policy
Advocacy leadership	A strong NGO community promoting ICTs in schools	
Gender equity		Inequality in access to education between boys and girls
Infrastructure and access		Erratic supply of electricity/High costs of telephone connection and the long distance charges

Collaborating mechanisms	A strong donor support for ICTs and government commitment to rebuilding education infrastructure	
Human resource capacity	Increase in private initiatives providing ICT training	Inadequate supply of skilled ICT labour in Sierra Leone
Fiscal resources		Lack of adequate government resource for education
Learning content		Lack of any standardised ICT curricula
Attitudes	Strong commitment on the part of teachers and administrators	
Sustainability		Heavy reliance on donor projects

Table 1: Factors influencing ICT

5 DHIS

The following sections provide a presentation of the development frameworks and coordination tools being used in the DHIS 2, as well as an overview of the design and key concepts of the system. Knowledge about these matters has been derived from my development participation within the project.

5.1 Development frameworks

In the following paragraph I will present the commonly used development frameworks in the DHIS 2 project. All of the mentioned frameworks are open-source and Java based.

5.1.1 Maven

Maven is a software project management and comprehension tool. Based on the concept of a project object model (POM), Maven can manage a project's build, reporting and documentation from a central piece of information and thus provide a uniform build system. Maven takes care of the dependency management by automatically downloading dependencies from a remote repository and installing them in a local repository, available to all projects. Maven can provide mailing lists and unit test reports, and offers guidelines to best practises to project directory layout and unit testing. DHIS 2 takes advantages of all of the mentioned features (Maven 2006).

5.1.2 Hibernate

Hibernate is an object-relational mapping system that let you store plain Java objects to a database. It is open source based and distributed under the GNU LGP License. Hibernate works only with relational databases, and only over JDBC. Hibernate's persistence strategy is known as transparent persistence because the model that are build contains no persistence code of any kind. Using Hibernate or similar systems have several advantages. The developer is able to work with objects instead of relational databases, and won't have to change the source code if another database is preferred. Hibernate generates SQL calls and provides automatic result set handling and object conversion. Hibernate was developed by a team of Java software developers around the world led by Gavin King (Hibernate 2006). DHIS 2 is currently being used with Resin, MySql and PostgreSQL as database management systems.

5.1.3 Spring

Spring is an open source lightweight application framework intended to make J2EE development easier. It consists of a container, abstraction layers for transaction management and JDBC, integration with ORM systems; aspect oriented programming functionality and a MVC web application framework. The layered architecture provides flexibility and the opportunity to freely pick the services needed. One main capacity of Spring is to wire application objects according to the principles of Inversion of control and Dependency injection. Spring lets the developer manage Java beans and dependencies in a complex system through a set of configuration files. All objects in DHIS 2 which provides services are mapped as beans, which keep the system easy to change and extend. Another main task of Spring is to promote good programming practise by enabling a POJO-based programming model (Spring 2006). DHIS 2 adheres to these guidelines by its extensive use of Java objects.

5.1.4 Struts and Velocity

Apache Struts is an open-source web application framework for developing Java EE web applications. Web applications differ from conventional websites in that web applications can create a dynamic response. Many websites deliver only static pages. A web application can interact with databases and business logic engines to customize a response. Struts uses and extends the Java Servlet API to encourage developers to adopt a model-view-controller (MVC) architecture. The framework provides three key components:

- A "request" handler provided by the application developer that is mapped to a standard Uniform Resource Identifier (URI).
- A "response" handler that transfers control to another resource which completes the response.
- A tag library that helps developers create interactive form-based applications with server pages.

The framework's architecture and tags are buzzword compliant. Struts works well with conventional REST applications and with nouveau technologies like Service Oriented Architecture Protocol (SOAP) and Asynchronous JavaScript and XML (AJAX) (Struts 2010).

Struts is used in combination with Velocity, which is a Java-based template engine. It permits the developer to use a simple and powerful template language to reference objects and variables defined in the Java code. Velocity provides separation of web design and code logic, making it more maintainable and fit for division of labour (Velocity 2006).

Every DHIS 2 web project uses Struts for the presentation layer.

5.1.5 JasperReports

JasperReports is a powerful open source Java reporting tool that has the ability to deliver rich content onto the screen, to the printer or into various formats like PDF, HTML and XLS. It is written in Java and can be used in a variety of Java enabled applications including J2EE or Web applications to generate dynamic content. Its main purpose is to help creating page oriented, ready to print documents in a simple and flexible manner. JasperReports provides the necessary features to generate dynamic reports, including data retrieval using JDBC (Java Database Connectivity), as well as support for parameters, expressions, variables, and groups. JasperReports also includes advanced features, such as custom data sources, scriptlets, and sub-reports. In the past, report generation has largely been the domain of large commercial products such as Crystal Reports. JasperReports is considered as the leading open source report engine and provides Java developers with a viable alternative to commercial software. JasperReports plays a central role in the DHIS 2 report tool module.

5.1.6 JFreeChart

JFreeChart is an open source free chart library that makes it easy for developers to display professional quality charts in their applications. JFreeChart's feature set includes a flexible design that is easy to extend and targets both server-side and client-side applications. JFreeChart allows you to easily incorporate advanced charting capabilities into Java applications, and has support for many output types including Java swing components, image files and PDF files. JFreeChart can be embedded and used to display graphs in JasperReports reports. The library is able to generate the most common chart types including pie, bar, line, and Gantt charts. JFreeChart is embedded in JasperReports and used in the DHIS 2 report module.

5.1.7 iReport

iReport is a powerful and easy-to-use visual report builder and designer based on the JasperReports report framework. iReport is a desktop application which is written in Java. It allows users to visually edit complex reports with charts, images and sub-reports. iReport is integrated with leading open source chart libraries for java such as JFreeChart. Report data can be retrieved in several ways including JDBC connections, Java beans, Hibernate and XML files. iReport is used to edit reports in connection with the DHIS 2 report module.

5.1.8 BIRT

BIRT is an Eclipse-based open source reporting system for web applications, especially those based on Java and J2EE. BIRT has two main components: a report designer based on Eclipse, and a runtime component that you can add to your app server. BIRT also offers a charting engine that lets you add charts to your own application. With BIRT one can add a rich variety of reports to the application including lists, charts, cross-tabs and documents. BIRT makes it possible to add totals, averages and other summaries of numeric data used in the report.

5.1.9 xStream

XStream is a simple library to serialize objects to XML and back again. It provides a high level facade that simplifies common use cases. It allows for serializing most objects without need for specifying mappings. Speed and low memory footprint are a crucial part of the design, making it suitable for large object graphs or systems with high message throughput. No information is duplicated that can be obtained via reflection. This results in XML that is easier to read for humans and more compact than native Java serialization.

5.1.10 Junit

JUnit is a unit testing framework for the Java programming language. JUnit has been important in the development of test-driven development, and comprises a family of unit testing frameworks collectively known as xUnit that originated with Sunit (Wikipedia D 2010).

5.2 Coordination tools

In the following section I will present the tools used for coordination in the DHIS 2 project.

5.2.1 Bazaar

Bazaar is a version control system that helps you track project history over time and to collaborate easily with others. Whether you're a single developer, a co-located team or a community of developers scattered across the world, Bazaar scales and adapts to meet your needs. There's no need to choose between central and distributed version control tools: Bazaar directly supports many workflows with ease.

Bazaar supports many best practices including refactoring, pair programming, feature branching, peer reviews and pre-commit regression testing. With true rename tracking for files and directories, merging changes from others simply works more often.
It's easy to migrate many existing projects including full history. Unable to migrate just yet? Track projects managed in Subversion, Git and Mercurial using standard Bazaar clients. Friendly, fast and efficient, Bazaar is free to use, embed and extend to meet your needs. Bazaar is by October 2009 used by approximately 80 000 people around the world on more than 9 000 software projects (Bazaar 2010).

5.2.2 Launchpad

Launchpad is a web application and web site supporting software development, particularly that of free software. Launchpad is developed and maintained by Canonical Ltd. On 21 July 2009, the source code was released publicly under the GNU Affero General Public License (Canonical 2010). As of October 2010, the Launchpad repository hosts almost 15 000 projects (Launchpad 2009). The domain launchpad.net attracted 1 million visitors by August 2009 according to a Compete.com survey (Wikipedia C 2010). Launchpad provides:

- Bug tracking
- Code hosting using Bazaar
- Code reviews
- Ubuntu package building and hosting
- Translations
- Mailing lists
- Answer tracking and FAQs

- Specification tracking

5.2.3 Mailing lists

The DHIS 2 project uses three different mailing lists that are integrated in Launchpad. The developer list is used for issues related to the technical development of the system. The commit list includes e-mails that are automatically generated when someone commits code to the repository. The user list contains issues and questions regarding the usage of DHIS 2. The mailing lists are stored in web archives. This is favourable when needs for going back and referring to previous discussions and decisions emerge.

5.3 Key concepts

In the following section I make a brief explanation of the key concepts, objects and terms of the DHIS 2 system, and provide a diagram which displays the relations between the various objects.

5.3.1 Overview

DHIS 2 is written in Java and has a three-layer architecture. The presentation layer is web-based, and the system can be used online as well as stand-alone.

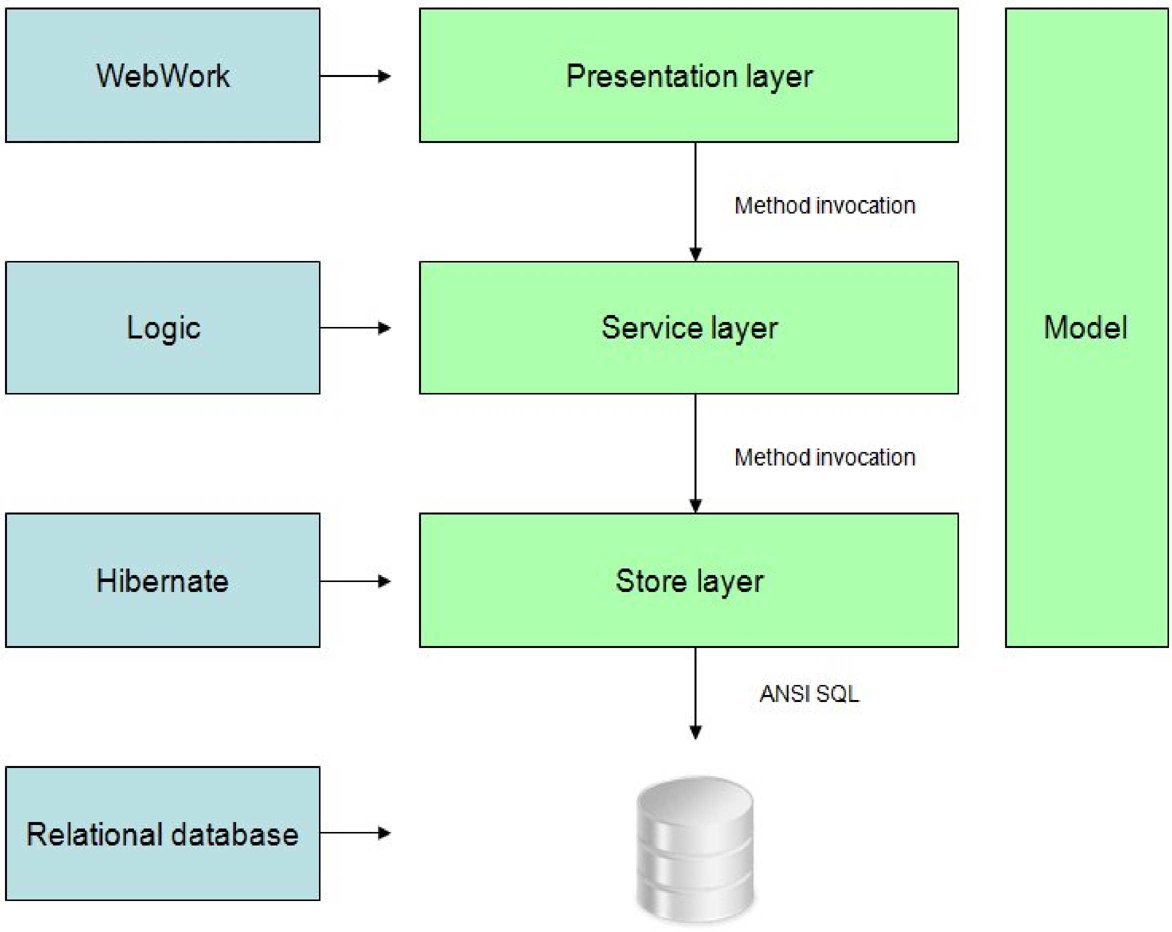


Figure 1: The layer structure of the DHIS 2.

5.3.2 Objects and data model

Data element

The data element object denotes a set of parameters related to a medical phenomenon like a diagnosis, treatment, procedure or physical actions performed by a patient, doctor or nurse. An example of a data element is Typhoid Fever - male under 1 year. Data elements can be grouped in data element groups.

Organisation unit

An organisation unit is the definition of any medical institution or statistical office at any level in the health hierarchy. An organisation unit can be a hospital, ward, district office, provincial office or the national ministry of health. The organisation units are organised in a tree hierarchy, implying that units may have a parent and a set of children. Organisation units can be grouped in organisation units groups.

Data set

The data set is a collection of data elements for which an organisation unit is supposed to collect and register data. The data set defines the time interval which the data is supposed to be registered for and the type of data supposed to be registered. A data set is tied to an organisation unit and may inherit from a parent data set. An example of a data set name is Notifiable Diseases Weekly.

Indicator

An indicator contains a formula which is intended to describe the state of a medical phenomenon. The state can be described as a rate or a ratio. An example of an indicator describing a rate is:

Infant tested for HIV = Infant tested for HIV / Live birth to woman with HIV

Data value

The data value represents a registered value in the database, and is identified by the data element and the period it was registered for and the organisation unit which registered the value. The registered value can be a number, a text sequence or a true-false value.

The data model is flexible in all dimensions in order to allow for capture of any item of data. The model is based on the notion of a DataValue. A DataValue can be captured for any DataElement (which represents the captured item, occurrence or phenomena), Period (which represents the time dimension), and Source (which represents the space dimension, i.e. an organisational unit in a hierarchy).

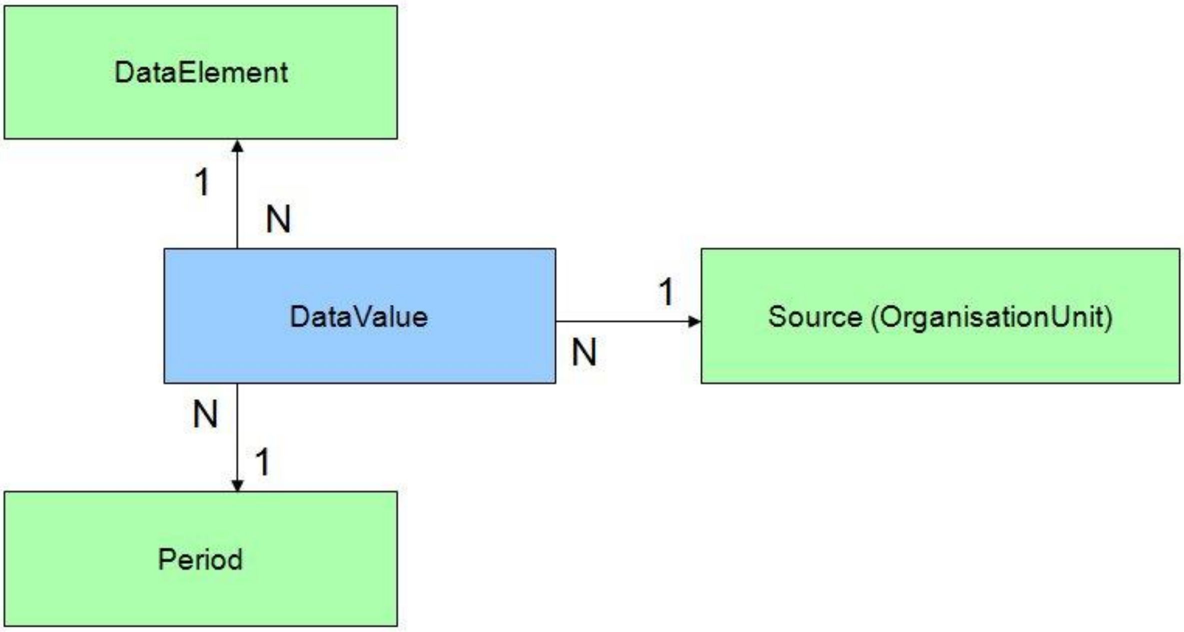


Figure 2: DataValue diagram

A central concept for data capture is the DataSet. The DataSet is a collection of DataElements for which there is entered data presented as a list, a grid and a custom designed form. A DataSet is associated with a PeriodType, which represents the frequency of data capture.

A central concept for data analysis and reporting is the Indicator. An Indicator is basically a mathematical formula consisting of DataElements and numbers. An Indicator is associated with an IndicatorType, which indicates the factor of which the output should be multiplied with. A typical IndicatorType is percentage, which means the output should be multiplied by 100. The formula is split into a numerator and denominator.

Most objects have corresponding group objects, which are intended to improve and enhance data analysis. The data model source code can be found in the API project and could be explored in entirety there. A selection of the most important objects can be view in the diagram below.

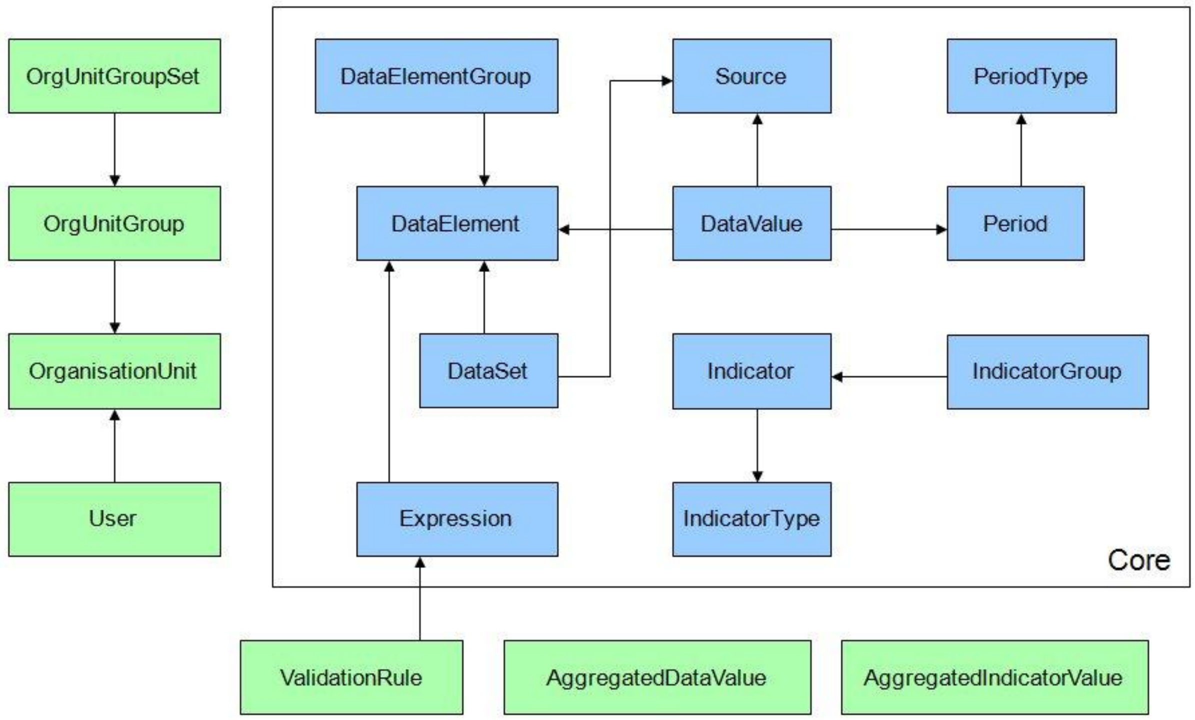


Figure 3: Class diagram showing the relationship between the core value objects

5.3.3 System design

The following section will give a brief explanation of the design of the DHIS 2 system. The DHIS 2 core modules are responsible for data persistence. Basic operations like retrieving, updating, adding and deleting data are performed by the store modules. The service modules contain the business logic, like functionality for aggregation of data, data mart, import and export, validation, user administration, user options, and internationalisation. Aggregation of data denotes aggregation over several periods of time and several organisation units in the hierarchy. Data mart is a kind of a data warehouse, where data is aggregated and exported to dedicated tables in the database. The advantage of keeping a data mart is related to integration of external information processors and reformatting of data. The DHIS 2 core also provides modules that support the use of application frameworks like Hibernate and Spring, testing, and transaction management.

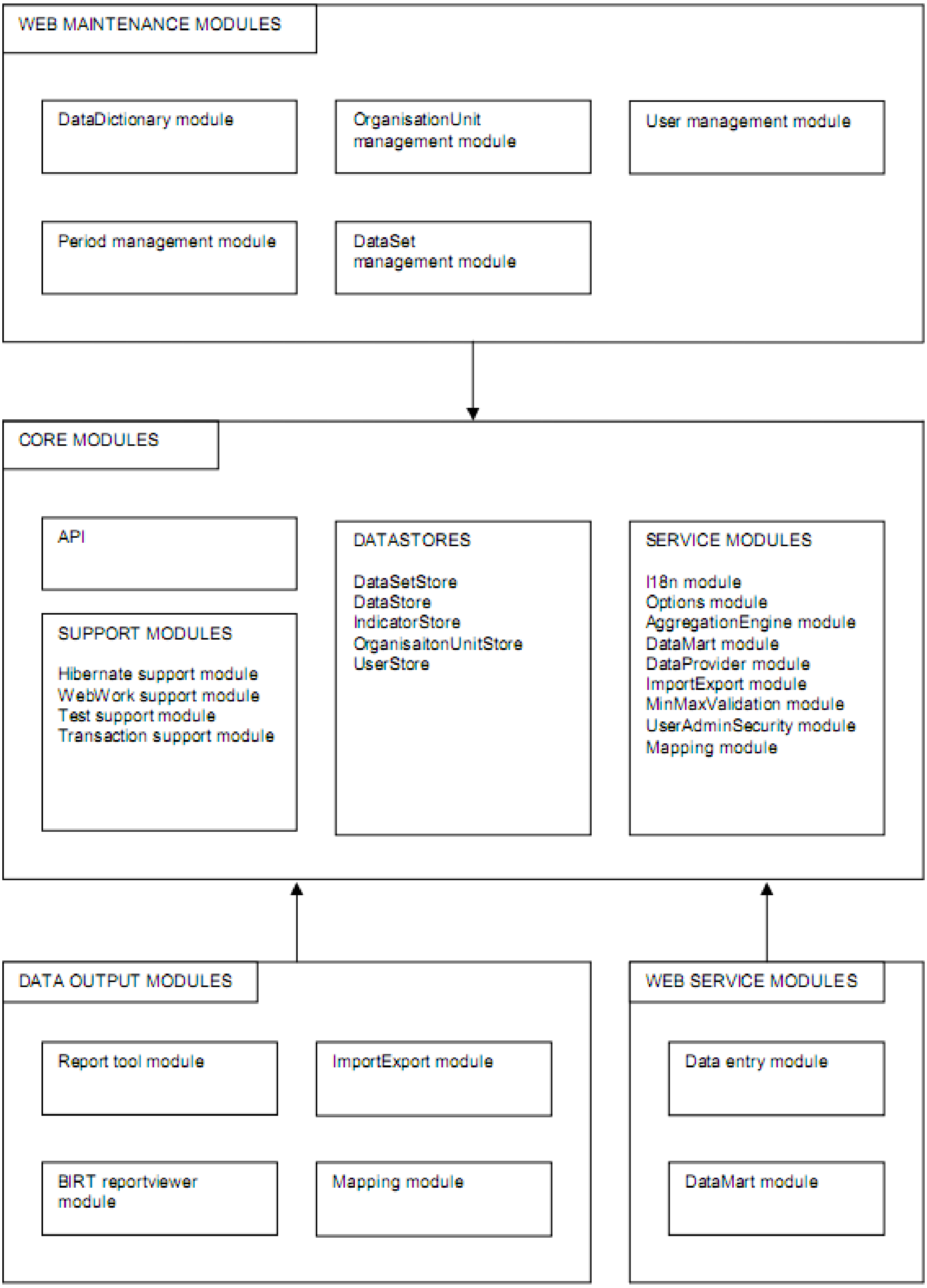


Figure 4: Overview of the modules in DHIS 2

DHIS 2 contains several web modules built upon the core. The web modules can roughly be divided in three. Firstly, the maintenance section provides administration of metadata definitions. Metadata in this connection refers to data elements, indicators, organisation units, users, periods, and data sets. These modules provide functionality for adding, deleting, updating and retrieving such data, as well as management of the respective groups.

Secondly, the data output section contains two reporting modules and an import-export module. The first report module is a general module optimized for creating detailed reports, and will act as basis for the discussion in chapter 8. The second report module is based on the BIRT reporting framework (ref. 5.1.8), which is favourable for dynamic and flexible reports. The import-export module uses an XML formatted file for interchange of data between DHIS 2 installations, and provides interfaces for complete and detailed exports.

Thirdly, the service section contains modules for data entry and datamart functionality. The data entry module lets the user enters data for a given organisation unit, data set and period. The datamart module provides the user with a basic and a detailed interface for exporting data to the datamart.

As demonstrated in Figure 3, the DHIS 2 design is highly modularised, and the presentation layer is composed of fairly independent modules. The design is intended to reflect the structure of the DHIS 2 project, which consists of development nodes in Oslo, Vietnam and India. Hence, modularisation is applied to the design in order to allow for distributed Development (Øverland 2006).

5.4 DHIS 2 Live

The DHIS 2 is also available as a light-weight distribution called DHIS 2 Live. This package is extremely convenient and easy to install as it contains everything you need in order to run DHIS 2 and is simply distributed as a compressed file archive. It is based on an embedded Jetty servlet container and an embedded H2 database. Jetty is a 100% pure Java based HTTP server and servlet container. It is a free and open source project under the Apache 2.0 Licence (Wikipedia G 2010). H2 is a Java based SQL database. It is very fast, open source and has a Java database connectivity (JDBC) API. Other features are:

embedded and server modes; in-memory databases, browser based console application and it has a small footprint, around 1 MB jar file size (H2 2010). To start up DHIS 2 Live you can simply unpack the archive and run the executable file.

5.5 Licence

DHIS 2 is open source/free software released under the BSD license and can be used at no cost.

6 The GIS module

The following sections provide a presentation of the development frameworks and coordination tools being used in the project, as well as an overview of the design and key concepts of the system. Knowledge about these matters has been derived from my GIS work over the last year.

6.1 Development frameworks

In the following paragraph I will present the commonly used development frameworks in the GIS module. All of the mentioned frameworks are open-source and based on JavaScript.

6.1.1 MapFish

MapFish is a flexible and complete framework for building rich web-mapping applications. It emphasizes high productivity, and high-quality development. It is based on the Pylons Python web framework. MapFish extends Pylons with geospatial-specific functionality. For example MapFish provides specific tools for creating web services that allows querying and editing geographic objects. MapFish also provides a complete ”rich internet application”-oriented (RIA) JavaScript toolbox, a JavaScript testing environment, and tools for compressing JavaScript code. The JavaScript toolbox is composed of the ExtJS, OpenLayers, GeoExt JavaScript toolkits, and specific components for interacting with MapFish web services. MapFish is compliant with the Open Geospatial Consortium standards. This is achieved through OpenLayers or GeoExt supporting several OGC norms, like WMS, WFS, WMC, KML, GML etc. MapFish is open source, and distributed under the GPLv3 license (MapFish 2010).

Only the MapFish client, not the MapFish server, is used in the GIS module for DHIS 2.

6.1.2 OpenLayers

OpenLayers makes it easy to put a dynamic map in any web page. It can display map tiles and markers loaded from any source. MetaCarta developed the initial version of OpenLayers and gave it to the public to further the use of geographic information of all kinds. OpenLayers is completely free, Open Source JavaScript, released under a BSD-style license. It is developed and supported by a number of organizations around the world. We are also looking for sponsors to help support the community. If you are in a position where you want to support the development of OpenLayers, but do not have development resources to share, you may be interested in supporting through our sponsorship program. OpenLayers is a pure JavaScript library for displaying map data in most modern web browsers, with no server-side dependencies. OpenLayers implements a (still-developing) JavaScript API for building rich web-based geographic applications, similar to the Google Maps and MSN Virtual Earth APIs, with one important difference - OpenLayers is Free Software, developed for and by the Open Source software community. Furthermore, OpenLayers implements industry-standard methods for geographic data access, such as the OpenGIS Consortium's Web Mapping Service (WMS) and Web Feature Service (WFS) protocols. Under the hood, OpenLayers is written in object-oriented JavaScript, using components from Prototype.js and the Rico library. The OpenLayers code base already has hundreds of unit tests. As a framework, OpenLayers is intended to separate map tools from map data so that all the tools can operate on all the data sources. This separation breaks the proprietary silos that earlier GIS revolutions have taught civilization to avoid. The mapping revolution on the public Web should benefit from the experience of history (OpenLayers 2010).

6.1.3 Ext JS

Ext is a JavaScript library for building interactive web applications using techniques such as Ajax, DHTML and DOM scripting. It was originally built as an add-on library extension of YUI, Ext includes interoperability with jQuery and Prototype. As of version 1.1, Ext retains no dependencies on external libraries, instead making their use optional.

Ext includes a comprehensive set of GUI-based form controls (or "widgets") for use within web applications that are able to communicate with web servers using AJAX:

- Text field and textarea input controls
- Date fields with a pop-up date-picker
- Numeric fields
- List box and comboboxes (also known as dropdownlists)
- Radio and checkbox controls
- Html editor control
- Grid control (with both read-only and edit modes, sortable data, lockable and draggable columns, and a variety of other features)

- Tree control
- Tab panels
- Toolbars
- Desktop-application-style menus
- Region panels to allow a form to be divided into multiple sub-sections sliders
- Charts

Ext includes web application support with features such as:

- Modal dialog boxes
- Interactive user-input validation prompts
- State management

Other features include a DOM selector class allowing operations to be performed on elements within the page, data stores that can be used to manage data, and classes to create and manage data in JSON and XML formats.

In April 2008 Ext 2.1 was released under a new dual license which allowed the options of the full GPL 3.0 license or a commercial license (Wikipedia E 2010).

6.1.4 GeoExt

GeoExt brings together the geospatial know how of OpenLayers with the user interface savvy of Ext JS to help you build powerful desktop style GIS apps on the web with JavaScript. It may be regarded as a geo-related extension to Ext JS as it offers a series of handy widgets for GIS. GeoExt is available under the BSD license and is supported by a growing community of individuals, businesses and organizations.

6.1.5 Struts and Velocity

The Struts framework and the Velocity template engine are used in the GIS server. See the DHIS chapter for a more detailed description.

6.2 Formats

This section explains what data formats that are read by the client.

6.2.1 *JavaScript Object Notation*

JavaScript Object Notation (JSON) is a lightweight data-interchange format. It is easy for humans to read and write. It is easy for machines to parse and generate. It is based on a subset of the JavaScript programming language. JSON is a text format that is completely language independent but uses conventions that are familiar to programmers of the C-family of languages, including C, C++, C#, Java, JavaScript, Perl, Python, and many others. These properties make JSON an ideal data-interchange language. JSON is built on two structures:

- A collection of name/value pairs. In various languages, this is realized as an *object*, record, struct, dictionary, hash table, keyed list, or associative array.
- An ordered list of values. In most languages, this is realized as an *array*, vector, list, or sequence.

These are universal data structures. Virtually all modern programming languages support them in one form or another. It makes sense that a data format that is interchangeable with programming languages also be based on these structures (JSON 2010).

The following example shows the JSON representation of an object that describes a person. The object has string fields for first name and last name, contains an object representing the person's address, and contains a list (an array) of phone number objects.

```
{
     "firstName": "John",
     "lastName": "Smith",
     "address":
     {
         "streetAddress": "21 2nd Street",
         "city": "New York",
         "state": "NY",
         "postalCode": 10021
     },
     "phoneNumbers": [
         { "type": "home", "number": "212 555-1234" },
         { "type": "fax", "number": "646 555-4567" }
     ],
     "newSubscription": false,
     "companyName": null
 }
```

A possible equivalent for the above in the widely known XML format could be:

```
<Person firstName="John" lastName="Smith">
  <address>
    <streetAddress>21 2nd Street</streetAddress>
    <city>New York</city>
    <state>NY</state>
    <postalCode>10021</postalCode>
  </address>
  <phoneNumber type="home">212 555-1234</phoneNumber>
  <phoneNumber type="fax">646 555-4567</phoneNumber>
  <newSubscription>false</newSubscription>
  <companyName />
</Person>
```

(Wikipedia H 2010).

6.2.2 GeoJSON

GeoJSON is a format for encoding a variety of geographic data structures. A GeoJSON object may represent a geometry, a feature, or a collection of features. GeoJSON supports the following geometry types: Point, LineString, Polygon, MultiPoint, MultiLineString, MultiPolygon, and GeometryCollection. Features in GeoJSON contain a geometry object and additional properties, and a feature collection represents a list of features (GeoJSON 2010). Below is a GeoJSON multipolygon sample with only one polygon (a district in Sierra Leone):

```
{
    "type":"FeatureCollection",
    "features":
    [
        {
            "type":"Feature",
            "id":"sl_districts_1",
            "geometry_name":"geom",
            "geometry":
            {
                "type":"Polygon",
                "coordinates":
                [[[-10,10],[-10,5],[-5,5],[-5,10]]]  // these coordinates will form a square
            },
            "properties":
            {
                "name":"Moyamba",
                "population_density":"29",
                "iso_country":"SL",
                < more properties >
            }
        }
        // add more feature objects
    ],
    "crs":
    {
        "type":"EPSG",
        "properties":
        {
            "code":"4326"
        }
    },
    "bbox":
    [
        -13.309010629848466,6.923379192855002,-10.270559353179449,9.999253303771024
    ]
}
```

The GeoJSON object has four main members. *Type* determines the type of the object and must be either "Point", "MultiPoint", "LineString", "MultiLineString", "Polygon", "MultiPolygon", "GeometryCollection", "Feature", or "FeatureCollection". In our case we want to have not only one polygon object, so *type* is set to “FeatureCollection”. The *features* array contains several feature objects which in turn contain a *type* member, a *geometry* object that defines the graphical figure and a *properties* object with information. The third root member, the *crs* object, specifies the coordinate reference system. The *bbox* array is the last root member and it includes information on the coordinate range.

6.2.3 Scalable Vector Graphics

Scalable Vector Graphics (SVG) is a language for describing two-dimensional graphics and graphical applications in XML. SVG images and their behaviours are defined in XML text files. This means that they can be searched, indexed, scripted and, if required, compressed. Since they are XML files, SVG images can be created and edited with any text editor, but specialized SVG-based drawing programs are also available. All major modern web browsers except Microsoft Internet Explorer support and render SVG markup directly. SVG allows three types of graphic objects: vector graphics, raster graphics and text. Graphical objects can be grouped, styled, transformed, and composited into previously rendered objects (SVG 2010).

6.3 Key concepts

In the following section I make a brief explanation of the key concepts, objects and terms of the GIS module of DHIS 2.

6.3.1 Overview

The GIS server is Java based and integrated into the DHIS 2 core which gives it a three layer architecture:
Communication with the database through Hibernate takes place in the store layer.
Objects, service functionality (logic) and an application programming interface (API) are implemented in the service layer.
Action classes are implemented in the service layer. These classes receive requests from the client, calls the appropriate methods in the service layer which provide the desired data back via the store layer. The actions classes then return the data to the client formated as JSON through Velocity templates.

The client is made up of pure JavaScript and the graphical user interface is implemented with Ext JS. The center component, the map area of the view port, are powered by OpenLayers and uses MapFish functionality to calculate and display geo-statistical information on top of it. GeoExt widgets are used for special geo-related tasks such as the advanced container based map layer tree.

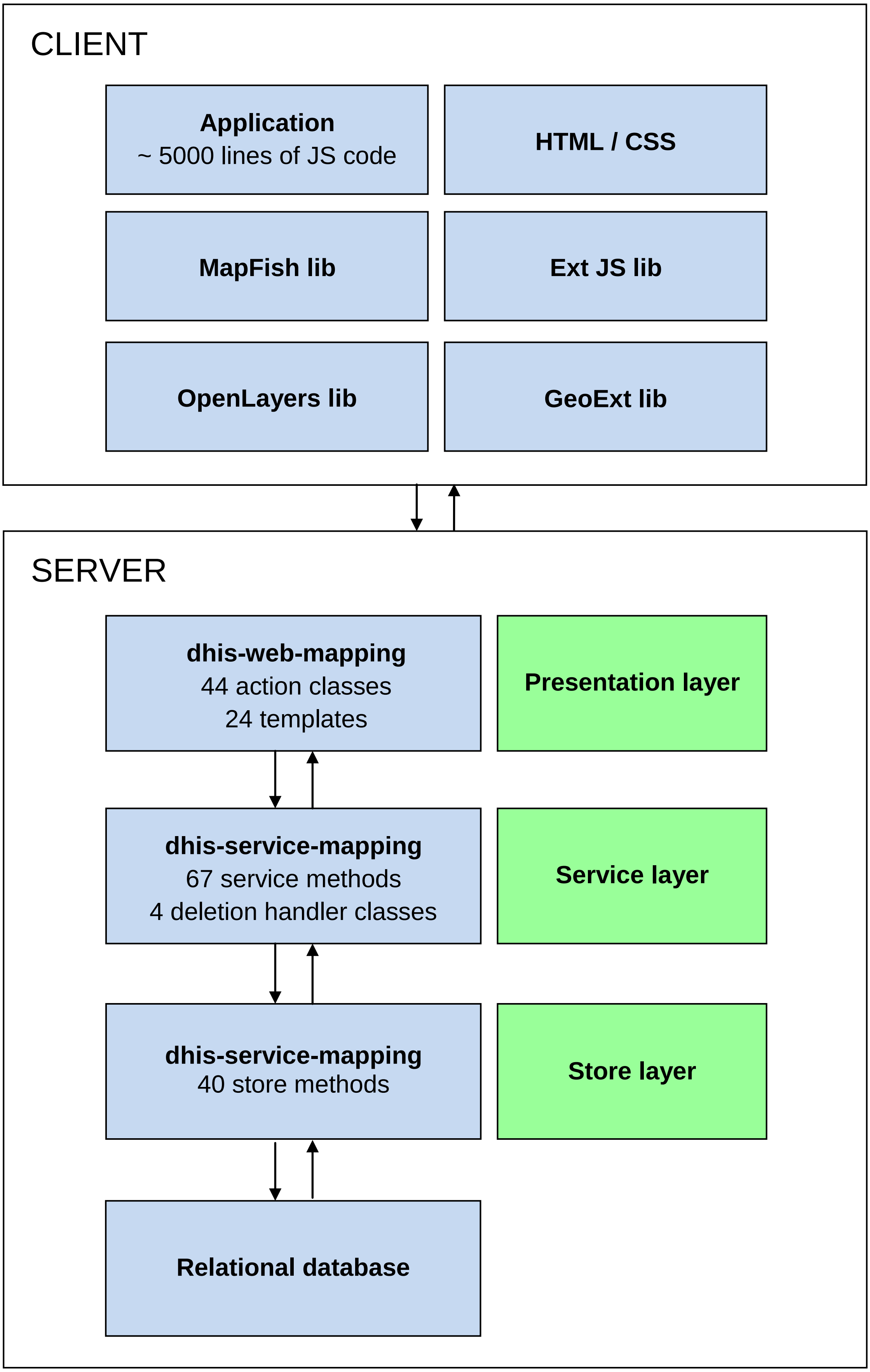


Figure 5: GIS module overview

6.3.2 Objects

This section presents the current GIS objects.

Map

The map object represents a map registered by the user. Every map has a superior and a secondary organisation unit level. The superior level defines the scope. This is usually set to national level because in most cases you would want to display the entire country. However, it could be set to e.g. province level if you have a particularly detailed map for one specific province and want to leave out the others. The secondary level defines what level you want to display data for. Thus, a map that is supposed to display data for e.g. all chiefdoms in Sierra Leone, the superior level should be set to national and the secondary to chiefdom.

In order to transfer, calculate and display data indicators for organisation units at the secondary level, the map object also holds the name the data column that should be matched against the database. In addition to tedious properties like name, longitude, latitude, zoom level and feature type (point/polygon/line), a specific *map source type* is persisted in the map object. Map source type is explained in the next section.

MapLayer

A MapLayer object stores information about a map layer that is added to the layer tree, either as a base layer (background) or as an overlay on the thematic map. An overlay needs graphical attributes like fill color, fill opacity, stroke color and stroke width. The fact that a map layer is made up by geospatial vector data similarly to a map, forces them to be grouped by the map source type as well.

MapLegendSet

A legend set is created by first defining a color for the lowest and the highest aggregated indicator values. A desired number of class breaks then splits up the legend in order to make the thematic map meaningful. The selected calculation method decides the size and how the different values (organisation units) are distributed within the prescribed number of classes. There are three different calculation methods:

Equal intervals, which is rather self-explanatory. The legend is simply split into intervals of the same size, like (*highest value* – *lowest value*) / *number of classes*.

Distributed values. This method tries to populate the classes equally, which may lead to heavily unequal intervals depending on the value dissipation.
Fixed bounds, which lets you define your own class breaks. Handy if you e.g. want to separate the organisation units above and below a single crucial value.
Finally you are supposed to link the legend set to one or many indicators. The point of doing this is that now the legend set settings will automatically be applied to the thematic map selections for the assigned indicators. Apart from the practical advantage this could be quite helpful due to the fact that for some indicators a high value is considered a good result and for others it might be a bad result. A typical scenario that might occur if you do not define a legend set for you indicators, could be when you change the indicator in your selection and forget to change the colors and your map presents the result upside down.

MapOrganisationUnitRelation
An organisation unit relation is related to a specific map object and defines a connection between an organisation unit in the database and an organisation unit in the shapefile. These relations are created by the administrator after a new map has been registered. The fact that the relations are stored means that this process needs to be done only once. The point is to let the application know which organisation unit an aggregated indicator value should be transfered to. Now that we have a genuine link between the respective organisation units we are able to speed up the process of transfering values when drawing thematic maps. This is because the organisation units do not have to be searched for and matched on their names every time.

MapView
The MapView object stores a thematic map selection, more exactly *indicator group, indicator, period type, period, map, method, classes, low color* and *high color*. This gives you quick and easy access to your favourite selections. You also have the possibility to add your favourites to the DHIS 2 dashboard. Thus, directly from the dashboard, you can start up the GIS application and display one of your favourite views with only one click.

6.3.3 Map source type

I have extended the system to support three different ways of feeding the application with geospatial vector data. The point is to avoid restricting the user to one single way of doing this and the different alternatives have different pros and cons, which I discuss in the analysis chapter. In the administrator panel you may set the *map source type* to either *GeoJSON, Shapefile* or *Database*. The selected map source type neatly decides the application's behaviour behind the scenes.

- *GeoJSON*
  The map registration panel will look for GeoJSON files that are placed in a specific subfolder of the DHIS2_HOME environment variable path. Only maps, map layers and map views with *geojson* as map source type will be available.
- *Shapefile*
  Gives you the opportunity to connect your private geospatial, server such as Geoserver, and load all your installed shapefiles directly into the application. Maps, map layers and map views will be stored with this map source type to separate them from objects of a different type. You might as well connect to a remote server.
- *Database*
  The shapefile generator described in the next section will be used.

6.4 Shapefile generator

The purpose of creating this shapefile generator, in short, is that I want people to be able to use the GIS application even if they do not have any shapefiles. If they, on the other hand, have the coordinates of the organisation units they want to display and analyse in the GIS, the shapefile generator makes this possible.

I will not give any technical presentation as it is not relevant for the discussion, but rather briefly describe how it works. The DHIS 2 offers functionality that lets you create new organisation units. I have extended these objects to store coordinates as well. When a new health clinic is opened somewhere in Sierra Leone, it can easily be registered in DHIS 2 with its coordinates. Now, by setting the *map source type* described in the previous section to *Database*, the shapefile generator collects all organisation units at the desired level and transforms the data into geospatial vector data in order to make it GIS usable. Then it is transferred to the client, rendered to one of the data analysis widgets I have created with MapFish, calculated and finally applied as a thematic map. The user of this thematic map is simply asked to select an organisation unit level instead of a registered map.

6.5 Client side code samples

To give an impression of how neat and easy it is to work with these JavaScript frameworks I want to present some small samples. See the developer documentation appendix for a more detailed description and a better overview.

Ext JS offers high user interface ability, but for particular releases you may not be entirely satisfied with the look or behavior of the controls. Then, I want to accentuate that it is remarkably easy to modify looks and behavior by editing the library. An even better idea, though, is that instead of making changes directly to the source code, you can alter these things by writing an override and execute it from the code during the initial page load. The advantage is that this way your modifications are not lost when the library is upgraded. Let us see how we can change the behavior of e.g. the layout. Ext JS offers several types of layouts, including *absolute, accordion, anchor, border, card, column, fit, form, table, vBox* and *hBox*, that may be used largely or to a limited extent. Among others I use *FormLayout* and I have modified its behavior and style from the code quite easily:

```
Ext.override(Ext.layout.FormLayout, {
        renderItem : function(c, position, target) {
            if (c && !c.rendered && c.isFormField && c.inputType != 'hidden') {
                var args = [
                    c.id, c.fieldLabel,
                    c.labelStyle || this.labelStyle || '',
                    this.elementStyle || '',
                    typeof c.labelSeparator == 'undefined' ? this.labelSeparator : c.labelSeparator,
                    (c.itemCls || this.container.itemCls || '') + (c.hideLabel ? ' x-hide-label' : ''),
                    c.clearCls || 'x-form-clear-left'
                ];
                if (typeof position == 'number') {
                    position = target.dom.childNodes[position] || null;
                }
                if (position) {
                    c.formItem = this.fieldTpl.insertBefore(position, args, true);
                }
                else {
                    c.formItem = this.fieldTpl.append(target, args, true);
                }
                c.actionMode = 'formItem';
                c.render('x-form-el-'+c.id);
                c.container = c.formItem;
                c.actionMode = 'container';
            }
            else {
                Ext.layout.FormLayout.superclass.renderItem.apply(this, arguments);
            }
      }
});
```

An OpenLayers example from the code could be the function where I add all overlays registered by the user to the map. A new vector layer is created with the desired protocol, strategy, name and the stored user settings. This function is executed both on the initial page load and when the map source type is changed in order to make the right overlays available to the map:

```
function addOverlaysToMap() {
	Ext.Ajax.request({
		url: path + 'getAllMapLayers' + type,
		method: 'GET',
		success: function(responseObject) {
			var mapLayers = Ext.util.JSON.decode(responseObject.responseText).mapLayers;

			for (var i = 0; i < mapLayers.length; i++) {
				var mapurl = MAPSOURCE == MAP_SOURCE_TYPE_GEOJSON ?
					path + 'getGeoJson.action?name=' + mapLayers[i].mapSource :
					path_geoserver + wfs + mapLayers[i].mapSource + output;
				var fillColor = mapLayers[i].fillColor;
				var fillOpacity = parseFloat(mapLayers[i].fillOpacity);
				var strokeColor = mapLayers[i].strokeColor;
				var strokeWidth = parseFloat(mapLayers[i].strokeWidth);

				var treeLayer = new OpenLayers.Layer.Vector(mapLayers[i].name, {
					'visibility': false,
					'styleMap': new OpenLayers.StyleMap({
						'default': new OpenLayers.Style(
							OpenLayers.Util.applyDefaults(
								{'fillColor': fillColor,
								 'fillOpacity': fillOpacity,
								 'strokeColor': strokeColor,
								 'strokeWidth': strokeWidth},
								OpenLayers.Feature.Vector.style['default']
							)
						)
					}),
					'strategies': [new OpenLayers.Strategy.Fixed()],
					'protocol': new OpenLayers.Protocol.HTTP({
						'url': mapurl,
						'format': new OpenLayers.Format.GeoJSON()
					})
				});

				treeLayer.events.register('loadstart', null, function() {
					MASK.msg = 'Loading...';
					MASK.show();
				});

				treeLayer.events.register('loadend', null, function() {
					MASK.hide();
				});

				MAP.addLayer(treeLayer);
			}
		},
		failure: function() {
			alert('Error: getAllMapLayers');
		}
	});
}
```

A simple example from the use of MapFish is the first step to creating a thematic map after a user has entered his desired selection in the user interface. This small piece of code is distributing the classified data to the point widget I created in order to display two different indicator values at the same time for health facilities; color and size (ref 8.4 for screenshots):

```
applyClassification: function(options) {
       this.updateOptions(options);

        var calculateRadius = OpenLayers.Function.bind(
                function(feature) {
                        var value = feature.attributes[this.sizeIndicator];
                        var size = (value - this.minVal) / (this.maxVal - this.minVal) *
                                (this.maxSize - this.minSize) + this.minSize;
                        return size;
                }, this
         );

        this.extendStyle(null,
             {'pointRadius': '${calculateRadius}'},
             {'calculateRadius': calculateRadius}
         );

         var boundsArray = this.classification.getBoundsArray();
         var rules = new Array(boundsArray.length - 1);

        for (var i = 0; i < boundsArray.length -1; i++) {
                var rule = new OpenLayers.Rule({
                        symbolizer: {fillColor: this.colorInterpolation[i].toHexString()},
                        filter: new OpenLayers.Filter.Comparison({
                                type: OpenLayers.Filter.Comparison.BETWEEN,
                                property: this.colorIndicator,
                                lowerBoundary: boundsArray[i],
                                upperBoundary: boundsArray[i + 1]
                        })
                });
                rules[i] = rule;
         }
         this.extendStyle(rules);

        mapfish.GeoStat.prototype.applyClassification.apply(this, arguments);
}
```

7 The GIS project

In this chapter I will explain and document the full development process behind the GIS application over the last year. Discussion and analysis follow in the next chapter.

7.1 Background

In this section I describe the relevant situation before I started working with GIS.

7.1.1 Personal context

The summer of 2008 I completed the required courses for my master's degree at the University of Oslo. With my master thesis in mind I accepted to join the HISP core developer team and started contributing to the DHIS. A year earlier I passed the course "Open Source Software Development", which is hosted by representatives from HISP team in Oslo. During this course I committed DHIS core source code, that is still in use, through a group assignment, and the three subsequent semesters I worked as a lab teacher for this course. Thus, I had some experience and knowledge of the DHIS before the autumn of 2008 when I was handed the responsibility of the new GIS project. On the other hand, I had no experience with GIS whatsoever. Also, the GIS team turned out to consist of me only so I was going to work alone, and there was hardly any technical knowledge on GIS within the HISP community at all. In other words, when it came to understanding GIS clients I had to start entirely from scratch.

7.1.2 The DHIS GIS context

HISP has seen several attempts to build a GIS for the DHIS fail over the last ten years. There are many reasons for this. One example is that they were all desktop applications, not web based, which brings along some huge limitations, the range of use taken into consideration. Two is the use of proprietary software that requires expensive licenses. Three is a limited and complicated integration with the DHIS. I discuss such limitations more thoroughly in the next chapter. In 2006 a new project called *OpenHealth* emerged. A lot of money was invested by the World Health Organization (WHO) in order to have a professional development team build a new monumental web based GIS solution and an OpenHealth integration module should make it able to run within the frames of DHIS. When I started to work with GIS in 2008, OpenHealth was nothing more than a simple half-finished prototype. It suffered from a serious lack of performance and was untouched for more than a year. There seemed to be no will at all to finally carry out the plan of having a useful web based GIS.

7.1.3 WHO, HMN and HealthMapper

There is a mutual interest of collaboration between HISP and WHO. The latter has realized that HISP is able to produce the software they want. On the other hand, HISP, which is in the middle of expansion, could use the resources and reputation of being the official software contributor of a world-famous organization like WHO.

WHO is the directing and coordinating authority for health within the United Nations system. It is responsible for providing leadership on global health matters, shaping the health research agenda, setting norms and standards, articulating evidence-based policy options, providing technical support to countries and monitoring and assessing health trends. The WHO's constitution states that its objective "is the attainment by all peoples of the highest possible level of health". Its major task is to combat disease, especially key infectious diseases, and to promote the general health of the people of the world. As well as coordinating international efforts to monitor outbreaks of infectious diseases, such as SARS, malaria, swine flu, and AIDS the WHO also sponsors programs to prevent and treat such diseases. The WHO supports the development and distribution of safe and effective vaccines, pharmaceutical diagnostics, and drugs. In addition to its work in eradicating disease, the WHO also carries out various health-related campaigns - for example, to boost the consumption of fruits and vegetables worldwide and to discourage tobacco use (Wikipedia I 2010).

Also, three former members of the Norwegian HISP team are now hired by the Health Metrics Network (HMN). This is a global partnership that facilitates better health information at country, regional and global levels. Hosted by the World Health Organization in Geneva, Switzerland, HMN aims to increase the availability and use of timely and accurate health information by catalysing the joint funding and development of core country health information systems. In pursuit of this goal, HMN lays out a vision and identifies strategies for health information system development and strengthening, supporting countries in implementing such strategies, and generating new knowledge and global public goods through research, technical innovation, and sharing lessons learned (Wikipedia J 2010). Of course, this makes the collaboration with WHO even closer.

HealthMapper is a desktop GIS application developed by WHO, described by their own website as a surveillance and mapping application that aims to address critical surveillance information needs across infectious disease programmes at national and global levels. It is a user-friendly data management and mapping system customized specifically for public health users. The system facilitates data standardization, collection and updating of data on epidemiology and on interventions and provides immediate visualization of data in the form of maps, tables and charts. Currently, it is or has been used by most countries in Africa, lots of countries in the east and far east, e.g. Ukraine, Kazakhstan, Turkey, Mongolia, India, Thailand, Indonesia, and also countries in America, e.g. Ecuador and Honduras. The ideas behind HealthMapper are to

- give the public health user a ready-made standardized digital database containing information considered essential by public health users including boundary maps, environmental factors (such as lakes, rivers, elevation) and vital information on basic population and basic health, school and water infrastructures.
- provide the public health user with a simple data management interface into which the user can easily enter and update public health indicators in a standard geographic format.
- provide the public health user with user-friendly icon-driven functions to automatically create maps, tables and charts of their data.
- make the system able to operate at both local and global levels.
- provide it to public health users at low or no cost.

HealthMapper has been designed and developed by WHO specifically for use by public health administrators working at national and district levels. It simplifies the collection, storage, updating, retrieval and analysis of public health data. It also simplifies the use of geographic information systems and mapping and provides a user-friendly interface to spatial analysis and data management (HealthMapper 2010).

The fact that WHO has chosen DHIS 2 as their new official health information system for developing countries creates a link between HealthMapper and the new GIS project for

DHIS. This means that the desired functionality from HealthMapper could be used as a template during the process of building the new web based GIS.

7.2 The initial work

This section describes the first two months of work on the GIS project.

7.2.1 Criteria

First of all, I had to be aware of the criteria and fundamental idea of this project. Simply put, the point was to have a geographic information system that was integrated with the existing system and could be easily distributed within the same software package. As a part of the DHIS 2, it was intended for, and thus had to be adapted to the context of, developing countries. The client had to be web based as opposed to the HealthMapper, which is a desktop application. To uphold the same licence standard it had to be open source and free software (FOSS) as well. This narrowed down the scope of development tools that could be used. To get fully integrated the server had to be included in the DHIS core to get access to the API and the necessary data. And the fact that the DHIS is database independent ruled out GIS specific database add-ons like PostGIS, MySQL Spatial etc.

The requirement specification was at first quite simple and composed by the HISP staff in Oslo. This included basic and fundamental functionality like loading indicator data from a DHIS database into the application, and finding a way to display the data in a map instead of using the current reporting tools.

7.2.2 Roadmap

To make sure I was going down the right path from the beginning with this project, I had to analyse the current situation and have a look at the realistic possibilities. There existed already a couple of half-finished web based GIS applications which could be interesting to learn from and possibly build upon. The first one I had a look at was the prototype mentioned in the DHIS GIS context section, called *OpenHealth*. I wanted to find out whether it was a good idea or not to take this solution further. The client was made up of JavaScript and offered a quite neat pivot table data selector. The server was a Mondrian online analytical processing (OLAP) server, written in Java. It supports the MDX (multidimensional expressions) query language and reads from SQL and other data sources. The available prototype had a couple of configured OLAP cubes. An OLAP cube is a structure that allows analysis of data and overcomes a limitation of relational databases; they are not well suited for near instantaneous analysis and display of large amounts of data (Wikipedia F 2010). Watching this application in action, however, did not whet the appetite at all as it was incredibly slow. It took up to 30 seconds to load, calculate and draw a very simple thematic map and the responsiveness of the map was really poor as well. Basic actions like panning the map back and forth required unacceptable amounts of time. There were among other things problems with query caching, no layer management at all, the dataflow was extremely inefficient, no select features, no legend management, the user could not save any favourite map views. I decided to leave this chunk of code alone. And the use of an OLAP server may give you advantages when you are dealing with truckloads of data, but for this purpose it was overkill and unnecessary hard to row.

I also wanted to have a look at another two existing prototypes. The first one was developed by a member of the HISP Vietnam developer team and had been tried out in India. The server was Java based and integrated in the DHIS core. Unfortunately, the implemented object model was not sufficiently thought-through, so it would have to be redesigned. The client was simply Struts, HTML and CSS which leaves way too many limitations when it comes to map operations and at least slightly advanced GIS functionality.

The second prototype was made on a hobby basis by a Tajikistani member of the Oslo team. An interesting open source Java library called Direct Web Remoting (DWR) was taking care of the communication between server and browser. DWR enables Java on the server and JavaScript in the browser to interact and call each other as simply as possible (Direct Web Remoting 2010). An OLAP server was again not well suited, though, and the client was not anywhere near rich enough to fulfil the GIS functionality requirements that would be expected of the new application. Also, these two prototypes were at a very early stage of development. I did not see any good reasons to utilize their source code and build upon them.

After I decided to start building a GIS module from scratch I did a lot of research to find the right tools. I presented my project, functionality requirements and criteria to GIS communities on the web and the response was overwhelming. Having the OpenLayers

JavaScript library as the cartographic engine seemed to be a mutual recommendation. The more specific GIS functionality could be developed by including the MapFish framework, which is made up of JavaScript as well. Regarding the graphical user interface (GUI) I was told that the Ext JS JavaScript framework both had a rich API and was rather good-looking. It also seemed possible to integrate in the DHIS portal solution so it appeared to offer what I needed. On top of everything, the GeoExt JavaScript framework was currently under heavy development. This project aimed to build geo-specific widgets made up by Ext JS and OpenLayers. It would not necessarily offer any features that I could use directly, but many of the developers behind the project were also part of the teams that developed MapFish and OpenLayers. Thus, on the respective mailing lists I could keep in touch with the developers that built the frameworks themselves, while we were simultaneously working with the same tools. This would hopefully provide the help that is needed when you are looking into four new comprehensive frameworks at once. The previous chapter offers more detailed descriptions of these frameworks.

The next weeks were spent trying to get to know these frameworks, understanding the API documentations to get as self-driven as possible and get up to speed on JavaScript in general. After playing with and expanding a couple of official samples and putting together my first application, I was convinced that I was able to develop something powerful with these tools. As I gained a better understanding of the technology I was working with, I started to perceive strengths and weaknesses as well. So far I had built a simple prototype that was able to display aggregated indicator data produced by the DHIS 2 data mart. An organisation unit and its indicator value were represented by a solid color that was applied on top of the actual area in the map. However, all the thematic map selections like indicator, period, map, and calculation method were hard-coded in configuration scripts that were read by the system. If I wanted to display a different map I had to change the code and reload the application. In other words, at this stage it was still not usable for others at all.

7.3 Professional training in Switzerland

In the beginning there usually turned out to exist simple solutions when I was stuck. Now, as I was working my way past the novice stage thanks to the rapid and effective mailing list interaction, I started running into more complex issues. Some of them did not lead to any response from the community at all. And I definitely did not know any other experts on the topic. At the same time, it was about time for me to produce an able prototype that I could present and have people to try out at the HISP workshop in New Delhi in March 2009. Also, there were interested people at the World Health Organization (WHO) that have followed the DHIS for a long time.

World Health Organization headquarters in Geneva, Switzerland

A minimum usability requirement that had to be implemented first was the ability to reload the thematic map widget with different geo-spatial vector data from the user interface. In practice, this means that the user could switch to a map at a different level and select different indicator data indefinitely without the need to reload the page, destroy the widget instance and create a new one, or change any hard-coded configuration. I was not able to figure this out myself at this stage and I definitely did not want to give up the tools that I just had started to enjoy.

In order to quickly progress out of the struggle I thought that making use of the HISP contacts at WHO could help me out. They had previously engaged the Swiss company that developed MapFish, called Camptocamp, which describes themselves as a service-oriented editor and integrator of open source software applications for geo-spatial solutions, business solutions, and infrastructure solutions. Based on the confirmed operational and prospective competencies and on the technical experience of its founding members in new information technologies, the company has been offering a complete set of services; consulting, research and development (R&D), training, and support since 2001. Their success is due to the continuous and dynamic contributions to various open source communities, to the extremely high technical expertise of its highly qualified and quality-driven professionals, and to the business partnerships fostered with its customers. Following a logic of durable development, Camptocamp commits to share all of its generic developments to the open source communities related to its business divisions (Camptocamp 2010). In February, WHO gave me the opportunity to meet developers from Camptocamp free of charge. During my stay in Switzerland I had two full days of personal training, the first one at the WHO headquarters in Geneva, the second in their own office at the University of Lausanne.

Camptocamp office at the University of Lausanne, Switzerland

This really gave me a moral boost and a better general understanding of how JavaScript applications are developed on the highest level. I absorbed valuable knowledge of MapFish, OpenLayers, Ext JS, and JavaScript in general, and we looked into more specific issues like the ones I had run into the previous weeks. I realized why I did not get any quick replies on the mailing lists. The issues were more complicated than first expected. During these two days we actually ended up doing some permanent enhancements to the MapFish core in order to make the functionality requirements of my application possible to implement. This was obviously a great experience for me and I was happy to see some mutual satisfaction as well, due to the useful improvements to their product. The rest of the month was spent implementing the requirement specification before the workshop in India.

7.4 HISP workshop in India

I decided to join this arrangement as I considered it a good opportunity to gain experience for the further development. The workshop was arranged in a city called Gourgon outside Delhi and was attended by members of the local HISP teams in Oslo, Delhi, Ho Chi Minh City (Vietnam) and Dar es Salaam (Tanzania). Here I would get in touch with people that have worked within the health sector of developing countries for a long time and are able to give me valuable insight and qualified opinions. I wanted to make the most of the two weeks I was supposed to stay in India and get as much input as possible. In order to avoid wasting time on a project that later turned out to be doomed to failure, my goal was to detect potential limitations and pitfalls at an early stage as well.

Workshop in Gourgon, India.

The fact that people generally are new to GIS definitely complicated the process of getting useful feedback. One of the first days of the workshop I handed out a questionnaire that hopefully would bring in valuable suggestions and information. People were supposed to draw on their knowledge from both the health sector and the IT sector in developing countries and present their thoughts on what they expect from a GIS for the DHIS, what kind of features they would find useful and so on.

Unfortunately, the majority considered themselves unable to contribute because they had no experience with GIS at all. And, the small response I got was just basic stuff that I had already contemplated and either discarded or found solutions to. I could not settle with the poor feedback and realized that I needed to engage people. To do this I

- gave an extensive presentation. Firstly, I demonstrated what the application could do so far from a user's point of view and people seemed to like it. I used maps and data from their own countries to increase the interest further. Secondly, taken into consideration that everyone had knowledge of the DHIS either as a user or as a developer, I explained how the GIS application is integrated into it. Lastly, for the developers that were interested, I presented the GIS client from a more technical perspective and even shared snippets of code.

- created and arranged two GIS assignments:
  - The first one was sort of a walk-through on how to get the application up and running in their own computers and thus mainly intended for users. I wanted people to get familiar with it and hopefully more interested in GIS in general. To speed up the process I handed out disks with the necessary files and software, as the available internet connection was not very fast.
  - The second assignment focused on the technical part. It requested those who could see themselves as future GIS developers to use the provided material to get their basic skills up to speed on modern JavaScript, get introduced to the different frameworks etc. Additionally, I offered personal assistance and training.

- invited everyone that did not want to set it up in their own computer to have a personal GIS lecture. This way, they could simply use my computer to try out the application and provide feedback.

These actions elicited a completely different enthusiasm than the clearly less motivating questionnaire. I was happy to see that people responded in a very positive manner. I mentioned that I wanted to discover potential limitations and pitfalls at an as early stage as possible. In this regard, the assignments turned out to be quite useful as two crucial issues were raised by those who installed and started playing with the GIS in their own computers. They were both related to the browsers' ability to handle huge amounts of geo-spatial vector data.

- The first one occurred when one of the developers tried to load a shapefile of his own country, converted to GeoJSON, into the application. This simply forced the browser to crash. After some debugging it turned out that there was nothing wrong with the file; the reason seemed to be the size of the script, which left me a rather terrifying feeling. Other web based GIS samples I had seen were mostly cities and other small areas. If country shapefiles were generally intended for desktop GISs and thus were going to crash your web browser, my whole web based GIS project idea for the DHIS 2 would be in danger.

- The second issue was related to the format GeoJSON as it seemed to be inefficient and suboptimized regarding the file size. What looked like unnecessarily many properties per object in order to validate as a vector data format was bad news as I had just discovered that modern web browsers were having a hard time handling those huge files. I decided to spend time and effort to figure these things out. As this is of great interest both for this project and for the web based GIS movement in general, I discuss this thoroughly in the next chapter.

In order to get the most of my stay in India, I decided to extend it with two weeks after the workshop had finished. The GIS implementor for HISP India wanted to scrap the prototype they had been trying out (described in the previous chapter) and start contributing to the new project. A good thing for me was that he got me in contact with representatives from the Ministry of Health. And even better, some of them had user experience with HealthMapper. I considered these two weeks very useful as I collected a whole lot of valuable feedback regarding GIS functionality and user interface from end-users. The next two months were spent building a more comprehensive prototype that actually could be used as a GIS application and not only a tool that could look promising in screen shots. I wanted to have something robust ready before June as I had the opportunity to join a one month stay in the West African country Sierra Leone, where three DHIS work shops were taking place.

7.5 Field study: Sierra Leone

Throughout this section I will mention the findings, in the shape of limitations and challenges, that I made note of during my stay in Sierra Leone. I will not report all the information I collected regarding user interface and most functionality requirements as I do not find them interesting to discuss in the next chapter.

7.5.1 Introduction

In India in March we were mostly staying in the hotel where the workshop was arranged or at the new and modern HISP office in Delhi. During this stay, on the other hand, I wanted to do a thorough field analysis. I wanted to see with my own eyes what the situation really was like in an African developing country and work with the people that actually are going to use the product. As mentioned in the Sierra Leonean context chapter, this is definitely one of the poorest countries in the world with a GDP per capita of around $900. The vast differences from my homeland, Norway, are well reflected through the Norwegian GDP per capita of around $60,000. The fact that Sierra Leone is in the middle of a recovery process after the civil war makes it even more interesting put in the terms of economy, health and information technology.

Even downtown in the capital of Sierra Leone, Freetown, the average standard of living is extremely low.

This stay in Africa turned out to be extremely useful for me. I was going to participate in three workshops crammed with true end-users. Some of the attendants had experience with both DHIS 2 and HealthMapper. Involving these people in the design process of the new GIS would be the optimal insurance in order to make the product usable and meet the users' needs. Moreover, I was going to visit the Statistics Sierra Leone (SSL) office in Freetown and explore their daily challenges and work routines. Also, we were going to have a dinner with the Ministry of Health during the stay.

7.5.2 Workshops

Besides me, the workshop arrangement team consisted of four health workers from the HMN. They were going to focus on training, discussion, and different user aspects of the DHIS 2, while I was responsible for the GIS part. The attendants were representatives from all the fourteen districts in Sierra Leone, where the DHIS 2 had been used for about a year.

The first workshop in Freetown attended by representatives from all the districts.

The first and the final workshop were arranged in the capital, Freetown, and the second in Bo, one of the centrally situated districts. They lasted for ten days altogether, and there were around 30 people attending each of them. Obviously, this was an ideal opportunity for me to gain understanding of the conditions out there in the districts and to get feedback on the work I had done so far. After having 5-20 minutes long interviews with at least ten persons per workshop, some issues, that were not related to GUI or functionality, were repeated over and over:

- Unavailability of up-to-date shapefiles.
- Inadequate funding available for training in new complex proprietary solutions, which results in a huge lack of qualified IT personnel. Their slight GIS experiences from the regional offices were bad as they found the current software very hard to learn.
- Inadequate funding available for acquiring new dedicated high-end hardware, which is required to run most proprietary solutions.
- No IT personnel can be dedicated to install, be trained on the use of, and support new complex proprietary software.
- Health professionals are over-committed and do not have time for training in GIS software operations with steep learning curves.

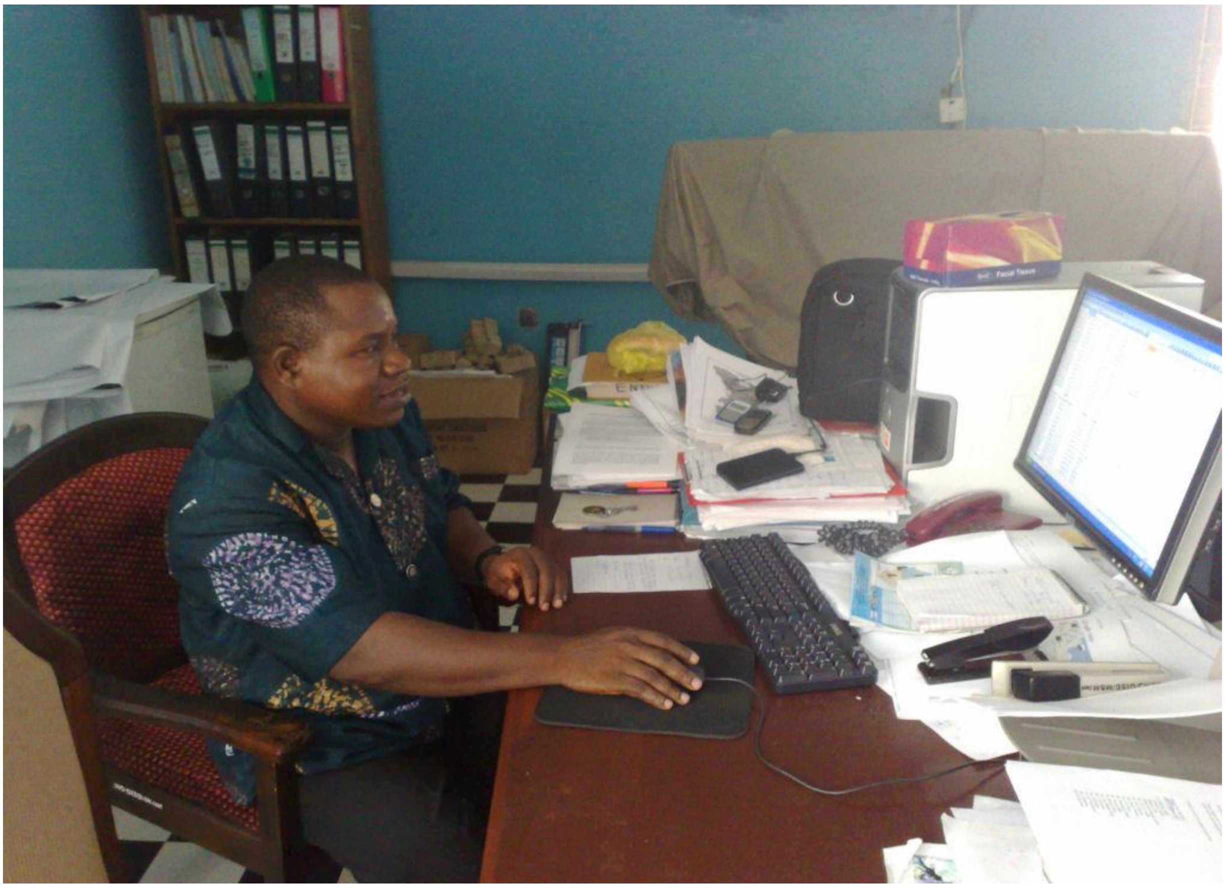
7.5.3 Statistics Sierra Leone

The Statistics Sierra Leone entrance.

Statistics Sierra Leone (SSL) was constitutionally effected by the 2002 Statistics and Census Act. SSL replaces the former Central Statistics Office, which had been in operation since independence in 1961. The mission of the Statistics Sierra Leone is to coordinate, collect, compile, analyze and disseminate high quality and objective official statistics to assist informed decision-making, and discussion within the government, business and the media, as well as the wider national and international community. To serve as the central authority for the collection, processing, analysis and dissemination of accurate, relevant, timely and high quality statistical information on social, demographic, economic and financial activities to serve the needs of users including the government and the general public. Statistics Sierra Leone also co-ordinates statistical activities and supervises the national statistical system.

The office has three main divisions. The *Economic Statistics* division focuses on providing robust economic indicators with the primary objectives of providing data for development planning, monitoring progress towards achieving the MDGs and Poverty Reduction, meeting the macroeconomic data requirements of multilateral agencies such as ECOWAS, IMF, World Bank, ILO, UNSD, etc. The *Demographic Social and Regional Statistics* division covers statistical data on population issues: education, health, gender and other social statistics. It provides timely and accurate data for policy planning and formulation for the country and it helps to design appropriate strategies to monitor the impact of some indicators on the labour force, poverty-reduction and economic growth within the country. The *Geographical Information System* is a computer-based information system that captures, stores, manipulates, analyses and displays spatially referenced and associated tabular attributed data. Geographic Information System has taken advantage of rapid development in micro processing technology to address the special challenges of storing and analysing spatial data (Statistics Sierra Leone 2010).

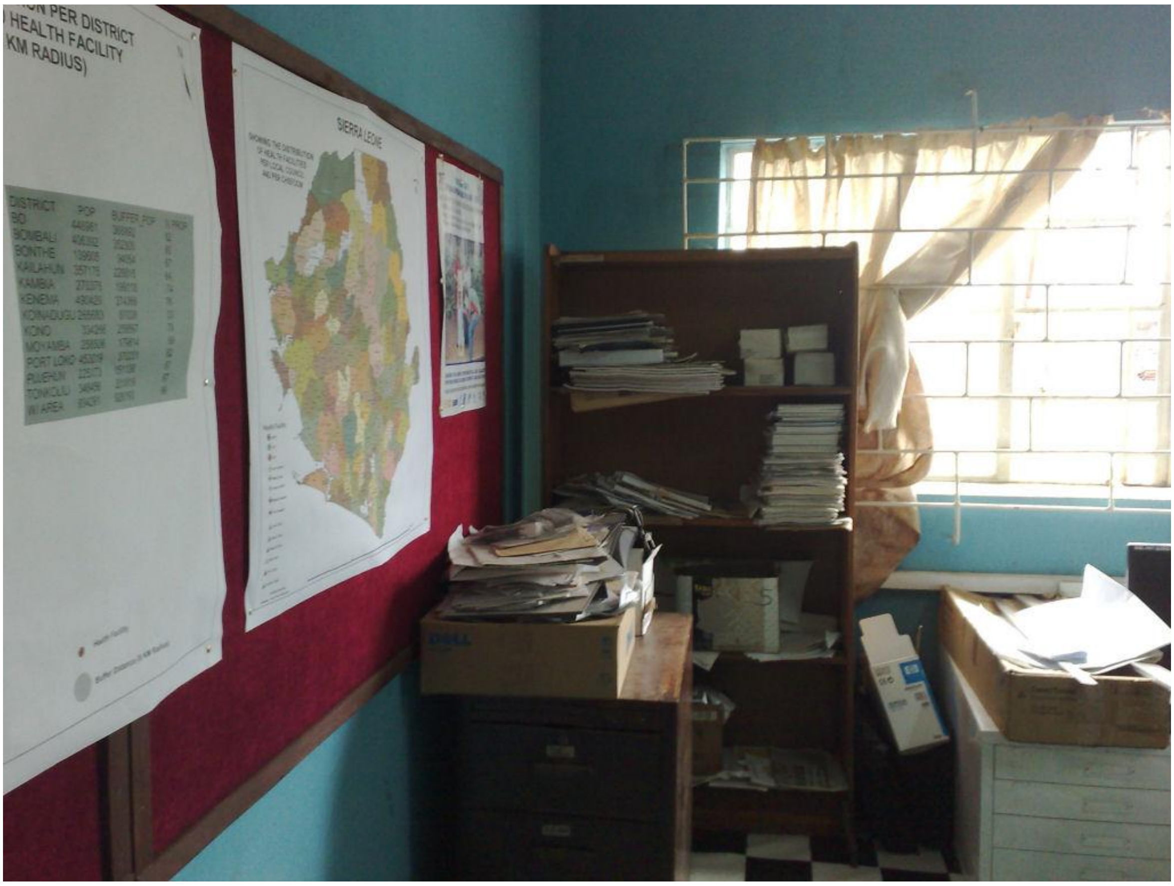
At the SSL I was welcomed by Senior Statistician Mr. Yambashu. He had kindly reserved several hours so we had the time to go through and analyse a regular work day of his.

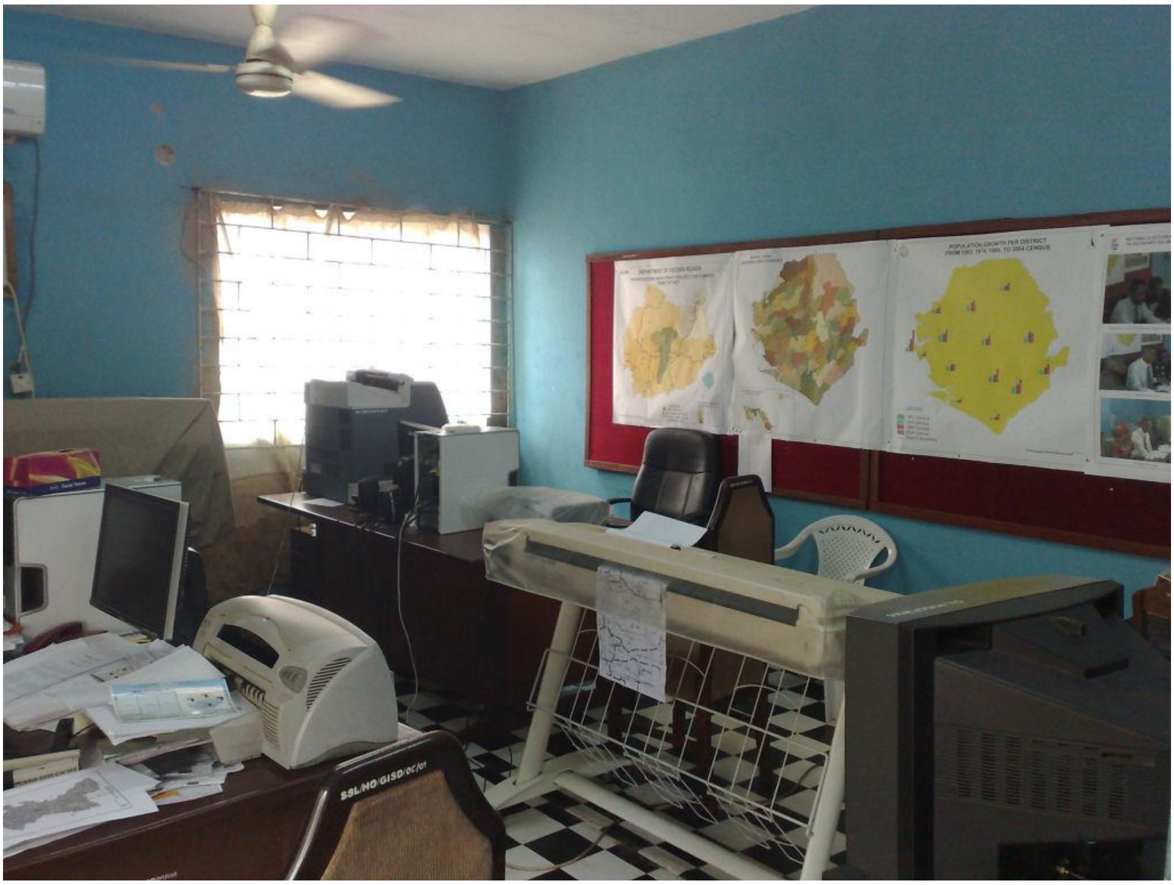


Senior Statistician Mr. Yambashu at the SSL office in Freetown.

During the meeting I made note of limitations and challenges that are relevant and interesting for my GIS project.

- Firstly, Mr. Yambashu showed me the GIS currently installed in his computer. This was proprietary software, which concerned him a lot as they had inadequate funding available for future software licences and maintenance fees.
- This software could probably have been useful here at the SSL, but he told me that it was barely used because their computer hardware could hardly deal with such heavy software. They were still practicing paper based map analyses. He demonstrated for me that performing spatial analysis with his computer was very slow, and his computer was forced to reboot during the presentation.





Paper based map analysis at the SSL.

- Public health users typically make use of only a small fraction of the functionality of large GIS packages.
- Public health databases are often in incoherent formats with no standardized geo-referencing, which makes their integration into a GIS difficult.
- The geographical information, sampling data, and the results of field studies are not easily and quickly accessible to public health officials to help them assess and respond to the situation.

After the presentation of his GIS in action I showed him the worth of my web based system. Mr. Yambashu was overwhelmed and expressed huge interest in the software. We were happy to see that it ran smoothly even on his slow computer. I explained that the application was still just a prototype, with insufficient functionality, but he told me that the thematic choropleth maps in particular were something he wanted to utilize.

I had the time to look around and talk to other workers at the office as well, and made some more interesting notes:

- Still, original data are kept in paper records. These have clearly limited distribution and are hard to access. Deterioration and loss are likely to occur.
- Results presented in conferences and journal articles do not provide easy access to the original, unfiltered data to current and future researchers.
- Limited distribution of spatial information and GIS know-how.

7.5.4 Ministry of Health

During the spring of 2008 an agreement with the Ministry of Health was found, and the DHIS 2 was rolled out and put to use in all districts in Sierra Leone. The last week of my stay, we got a dinner invitation from the Ministry in connection with the retirement of one of our officers from the HMN.

Dinner with the Ministry of Health.

This evening I had long GIS discussions with the representatives from the Ministry and I summed them up in the following notes:

- There is a need for an efficient, easy-accessible, and user-friendly geospatial information system without steep learning curves, which adequately addresses their specific needs.
- Agreement on new proprietary GIS solutions would demand significant resources and significant changes to existing plans and IT infrastructures. They expressed huge concern for dependency on proprietary foreign solutions.
- Need for technological solutions that do not require large financial, human, and technological resources.
- Need for technological solutions that are easy to integrate with existing IT infrastructure and capacity.
- Need for data and systems that have the capability of being interoperable.

A functionality requirement that I find interesting to discuss in the next chapter, and thus is mentioned here, was also presented by the representatives:

- Possibility to connect to an external repository that is running a GIS server. This way all GIS users can load shapefiles from this repository into their application, as opposed to their current GIS solution where everyone needs to have the shapefiles on their local hard drive.

Moreover, I talked to some of the representatives that had experience with HealthMapper. I made notes of what they thought were the best features and what needed to be improved, as it probably would be useful later on in the development process.

7.5.5 Unavailable shapefiles

One of the last days before I left I headed back to the SSL. The workshop interviews had told me that they could not get their hands on updated shapefiles, but in order to improve the quality of my GIS setup for Sierra Leone I wanted to try to get them from the SSL personally. The ones I was handed by HISP were a couple of years old and thus outdated. I was primarily after the newest shapefile with coordinates and information about the medical health care clinics. As mentioned in the Sierra Leone context chapter, such clinics are opened every week throughout the country after the civil war, which means that a facility shapefile from 2007 is virtually useless. The polygon shapefiles, though, do not contain as much population data and demographic information as the clinics shapefile, and the number of units (districts and chiefdoms) and their boundaries (coordinates) obviously do not change very often. In other words, it is not crucial that these are entirely up to date.

I was quite disappointed to leave the SSL office empty-handed. When I got back to the hotel I told my workshop partners that they would not give me what I wanted, but our most experienced companion from the HMN, Dr. Bob Pond, was not surprised. He is a public health physician that has been working for African countries for more than 20 years. Before working with HMN, he devoted his career to the development of health services in sub-Saharan Africa, being actively involved in the improvement of health services global monitoring and evaluation of the management of childhood illnesses in Ghana, Nigeria and Burkina Faso.

Later that night Dr. Pond gave us some background information regarding what happened to me at the SSL. I have summed it up in the following quotations:

"If you want to do a national scientific survey you need to select a sample that are going to be representative of the country. And to do that you need a distribution of the population. You need to have what the statistician calls a sampling frame, where you go down and scientifically select the location for the sample that you pick."

"The statistics office in Freetown uses their population data to select the scientific sample that is going to represent the country. You cannot do that unless you have access to demographic data and the spatial distribution of the population. And every time someone wants to do a national scientific survey they have to go to the statistics to get the necessary data. The statistics, though, will not give them that, even if they are qualified, unless they get paid for it. In other words, they sell their data and services to the highest bidder."

"If you ask the UN statistics office in New York, they will actually justify this. I say that this is maintaining corruption. This is information that is paid for publicly, but then it is used for private game. The statistical people refuse to say as a basic principle that this information should be released to qualified researchers, so that they can do their extra studies. There are so many examples of qualified researchers that are not able to get their hands on the necessary data."

”A justification for this is that it should be the statistical office that controls who does surveys. They have the legal responsibility to oversee the collection of statistics, put in terms of national ownership and national capacity. But the fact that you have to go and pay off the national statistics office in order to do statistical health research, though, has nothing but restrictive influence on the capacity. In a country like Sierra Leone, which is known as one of the poorest countries in the world, it is absolutely crucial to place money and effort where it is needed. Statistical health research will reveal the information you need to do this, e.g. what districts and chiefdoms have the highest infant mortality because of lacking vaccination, details on the state of the health facilities and so on.”

This gave me a clear picture of why I did not get the shapefiles I wanted from the SSL. What worried me the most was the fact that outdated shapefiles would definitely lower the quality of my (or any) GIS. In the next chapter I discuss this issue and the shapefile generator I implemented to solve it.

7.6 GIS meeting at WHO

During the summer WHO wanted to have HISP as their main software provider. Thus, in August, I went back to Geneva to discuss the future of the GIS project. I gave a presentation of my latest prototype to the people who are responsible for the software at WHO. They all seemed very pleased and interested in the further development. The HealthMapper experts announced what was the most important functionality that my GIS application still did not have, and thus they wanted to see replicated. These recommendations combined with the HealthMapper notes I made in Sierra Leone led to a long list of features that I am going to implement in 2010. Together, we registered the most urgent ones in Launchpad, and there are currently fifteen blueprints that are assigned to me:

- User interface: ability to add new WMS layers to the map
- Run data mart automatically when indicator, period and level are selected
- Organisation unit level drill down when a polygon is clicked
- Ability to display data elements (“raw data”) as well as indicators
- Filtering map extent: the map showing only a limited area such as a province or district
- Include fixed bounds in map views

- Graduated symbols for points
- Provide internationalization (i18n) of the user interface
- Ability to turn on and off organisation unit name labels
- Ability to search and locate an organisation unit in the map
- A filtering function for the organisation unit name list
- Organisation unit profiles as pop-ups which show values for all indicators in an indicator group
- Persist zoom and center level during organisation unit relationship assignment
- Ability to print map, legend and comments as pdf
- Display time series in the map

7.7 Summary

The following timeline sums up the major events, in which I participated, during the development process of the GIS module. The ones that are not boldfaced are not described in the empirical part of this thesis as I believe they are rather unimportant to the outcome of this study.

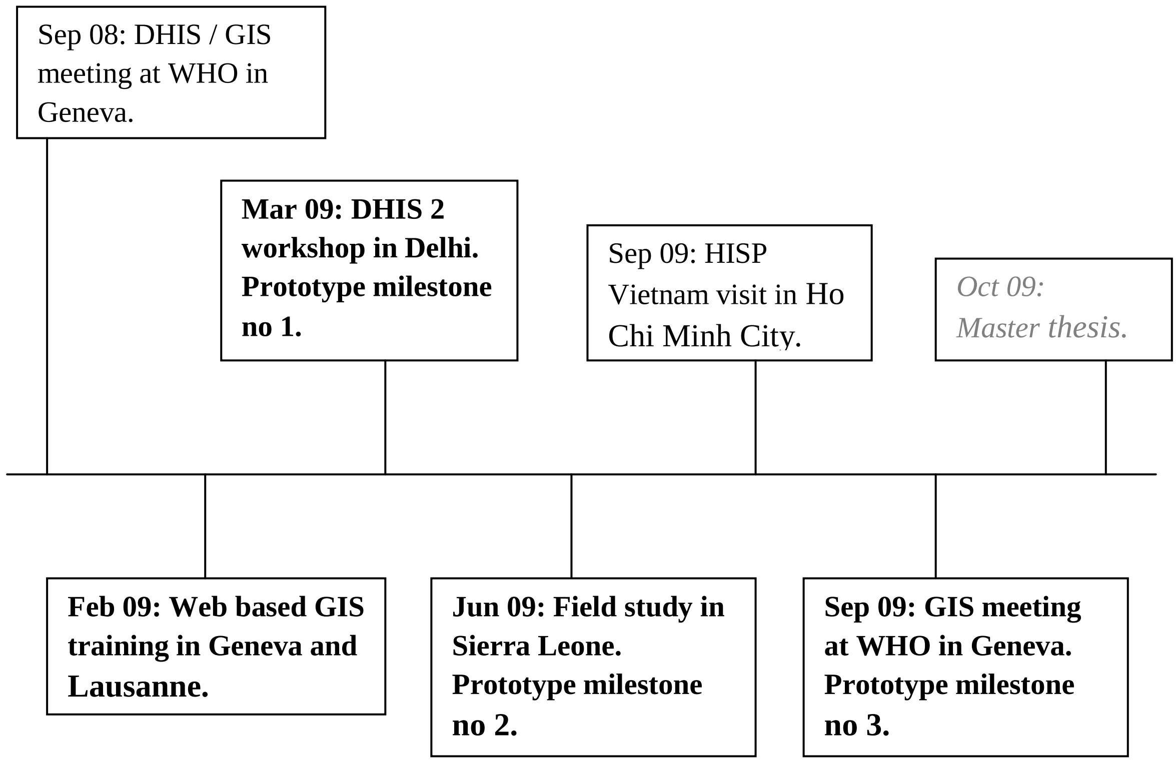


Figure 6: GIS project timeline

8 Discussion

In this chapter I will explore my research questions and discuss the empirical findings in relation to the presented literature.

Research objective:
Explore the capability of web based open source GIS development frameworks and open source methodology to improve sustainability of GIS implementations in developing countries.

To answer the research objective I have dispersed it into two questions:

Research question 1:
Explore the technical capability of web based FOSS GIS development frameworks.

Research question 2:
Explore limitations and solutions regarding development and implementation of GIS in developing countries in general, and whether web based FOSS GIS development frameworks and open source methodology can improve sustainability of such implementations.

To be able to make a qualified conclusion to the research objective I have connected the two research questions by personally utilizing the mentioned GIS technology to develop a GIS application that is based on elements from open source methodology and field research in Sierra Leone. In the following sections I discuss the research questions in relation to the evaluation of my own project.

8.1 Evaluation basis

In this section I account for the basis of the evaluation of my own project, which is conducted in the next section.

8.1.1 Defining and measuring success and failure

In order to evaluate my own project I will use Heeks' article (2002) about information systems and developing countries as a foundation to define success and failure. His work is based on qualitative review of a large number of case studies, e.g. Roche & Blaine (1996), Odedra-Straub (1996) and Avgerou & Walsham (2000). He is aware that categorization (success or failure) runs into some immediate difficulties that are hard to completely resolve; subjectivity and timing of evaluation. Viewed from different perspectives, according to Lyytinen & Hirschheim (1987) and Sauer (1993), one person's failure may be another's success. Still, Heeks' categorization tries to address this within the limits imposed by the subjectivity of the case study writers themselves. This combined with the fact that the author, working at the Institute for Development Policy and Management at the University of Manchester, does not seem to have vested interests in the topic, makes me trust and place emphasis on this article.

The first category, *total failure*, is described as an initiative that was never implemented or abandoned immediately after the implementation. The second, *partial failure*, is a system that was implemented, but were major goals are unattained or there are significant undesirable outcomes. The third and final, *success*, is an implementation where most stakeholder groups attain their major goals and do not experience significant undesirable outcomes (Heeks 2002).

8.1.2 The extent of success and failure

Heeks emphasizes that no one knows for certain what proportion of developing country IS projects fall into each of the three outcome categories. The question is hard enough to answer in the industrialized countries, where there are, at least, a certain level of surveys, evaluations, and analysis present (Korac-Boisvert & Kouzmin 1995, James 1997, Sauer 1999, The Economist 2000). On the basis of the range of figures provided in these surveys, one may estimate that something like 1/5 to 1/4 to one-quarter of industrialized country IS projects fall into the total failure category, about 1/3 to 3/5 fall into the partial failure category, and only a minority fall into the success category (Heeks

2002). Heeks claims that despite the weak evidences, it all points in one direction; toward high rates of IS failure in developing countries (ref 2.2.5).

8.1.3 Design-actuality gaps

Now, since there are high rates of IS failure in developing countries, Heeks wants to know why and seeks to understand developing countries better. He focuses on what he calls a *design-actuality gap* which is described as the mismatch between local actuality (where we are now) and system design (where the design wants to get us). In practice, because of subjective expectations about the future and subjective perceptions of reality, it could be argued that every individual IS stakeholder has their own design and their own version of actuality. The following figure illustrates his design-actuality conception.

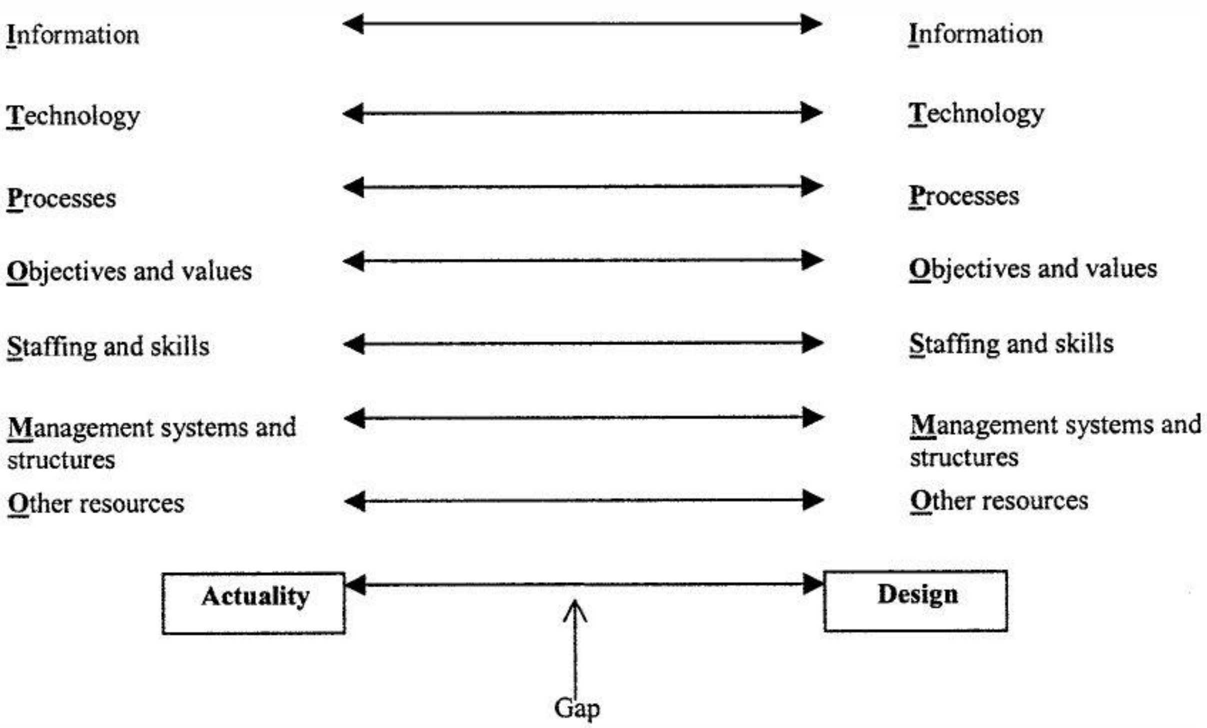


Figure 7: Design-actuality gap

Heeks says that the most extreme form of design-actuality gaps occurs when industrialized country designers create an information system within and for an industrialized country context, and that IS is subsequently transferred to a developing country. In such situations, the actuality of local conditions in the developing country will not have been considered at all in the original design, and a considerable design–

actuality gap is therefore likely, leading to a significant risk of IS failure. Sandiford et al (1992) states that the socio-economic realities and priorities of the third world are quite different (ref 2.3) and, if GIS is to be used for the challenges facing developing countries, then it must respond to those realities.

Heeks' further explanation is in my opinion particularly interesting and of great relevance to my own project. He claims that even if some effort is made to develop an information system specifically for a developing country, similar problems can arise. Industrialized-country stakeholders, such as consultants, IT vendors or aid donors, often dominate the IS design process in developing countries. Those stakeholders bring their context with them and inscribe it into their IS designs; inscriptions that will mismatch developing country actuality. Some stakeholders bring with them the "if it works for us, it will work for you" mentality that makes no attempt to differentiate between industrialized and developing contexts. Others will differentiate, but - given their poor understanding of local developing country conditions - their assumptions about user actuality will be incorrect. In all cases, large design–actuality gaps and high failure risks are the outcome (Heeks 2002). An example is the United States Agency for International Development (USAID), the United States federal agency responsible for administering civilian foreign aid. They deserve commendation for providing a great portion of the development aid in Africa, but at the same time they tend to support American companies, who often convey a "we know better" mentality, instead of local organizations (like HISP South Africa) with better knowledge of the actual context.

Both Al-Romaithi (1997) and Georgiadou et al (2005) are of the opinion that GIS technology is a product of the developed world (ref 2.2.6), and introducing such systems in developing countries involves large complexities dealing with social, technological, political and contextual issues. Taylor (1991) says that like most other information technologies, GIS is primarily a "first-world technology," an artifact of industrial and postindustrial societies in the developed world. Martin (1998) states that cultural differences in concepts of time, scale, detail, distance, values, topology and relationships mean that GIS implementation is context sensitive. Beyond these cultural differences, GIS implementation is also affected by institutional contexts and organizational interrelationships. Thus, the implementation of GIS in non-western settings requires a flexible and context sensitive approach, involving a variety of modifications to suit local needs. Successful investigations of GIS installations in non-western contexts require an approach that analyses the interactions between the technology and the specific social or institutional setting (Martin 1998). Many scholars have emphasized that system development approaches in developing countries have failed to consider context, and as a result projects tends to fail (Braa 1997, Walsham 2001). Sahay and Walsham (1997) point out that (ref 2.3.2) one of the main reasons for failure in GIS implementations is that very little attention is given to adapting the technology to the needs and capabilities of the countries and organizations in which it is going to be adapted. Taylor (1991) urges that to develop any local GIS system for micro level decision-making, it should be introduced, developed, modified and controlled by local users in their context. In order to (ref 2.3.1) achieve sustainability according to Braa et al (2004), the system must be shaped and adapted to the given context. These statements support the two elements compounded by Heeks (2002) to be the main causes of design-actuality gaps, and thus the main cause of IS failures in developing countries:

- *Components from the designers' own context*: IS design is a situated action - an action "taken in the context of particular, concrete circumstances" (Suchman 1987). This action draws elements of that context into the design: "Our technologies mirror our societies. They reproduce and embody the complex interplay of professional, technical, economic and political factors" (Bijker & Law 1992). Designers themselves are part of and shaped by that context, so their own cultural values, objectives, etc. will be found inscribed in the design (Shields & Servaes 1989, Braa & Hedberg 2002).
- *Conceived assumptions about the situation of the user*: This includes assumptions about the users' activities, skills, culture, and objectives, and assumptions about the user organization's structure, infrastructure, etc. (Boehm 1981, Suchman 1987, Clemons et al 1995, Wynn & deLyra 2000).

To summarize this section; in order to achieve success when introducing an information system in developing countries, the mentioned authors emphasize a *context sensitive approach* and *modifications to suit local needs* as vital. If these elements are disregarded you are most likely to extend the long list of IS failures in developing countries. This is because design-actuality gaps arise which means that the system's features will not meet the users' needs. I discuss this in relation to my own project in the next section.

8.2 GIS module evaluation

It is too early make a definite conclusion to whether my GIS project is a success or a failure. I say that because there is only five months since it was included into the official release of the DHIS 2. There is also a lot of functionality that is advertised for which is still not implemented. This requires further evaluation and new changes to the user interface. The warm reception is has received, however, indicates a rather bright future. Let us have a look at the development process and the literature combined.

8.2.1 Context sensitive approach

Al-Romaithi (1997), Martin (1998), Georgiadou et al (2005), Braa et al (2004), Walsham (2002) and Heeks (2002) all agree that building or adapting GIS/IS to developing countries imply some difficulties. As confirmed by Braa (1997), Walsham (2001), Suchman (1987), Bijker & Law (1992), Shields & Servaes (1989), Braa & Hedberg (2002), Boehm (1981), Clemons et al (1995) and Wynn & deLyra (2000), design-actuality gaps easily arise when developers design systems for a different context (ref 2.2.5).

Components from the designer's own context

In my case, a developer from the industrialized part of the world (me) has implemented a GIS, which should be considered (ref 2.2.6) first-world technology and a product of the developed world (Al-Romaithi 1997, Georgiadou et al 2005, Taylor 1991), for a context that I have shown is vastly different; developing countries. In other words, it is reasonable to say that this product has been vulnerable to design-actuality gaps during the process. On the other hand, the literature (ref 3.2.2) has told us that *participatory design* is a design approach that may counteract such gaps. Going step by step by making prototypes and involving users, by treating them as co-developers, in the development process are likely to ensure that the product meets the users' needs and is usable. By looking at my process in the previous chapter, we see that I have built a new prototype for three major milestones. The first one for India was based on basic initial requirements from the Oslo team in order to run it as a DHIS 2 module. The second one for Sierra Leone was expanded and rebuilt based on the qualified feedback I got in India. Then, the same process was reiterated before the meeting in Geneva based on the feedback from Africa. According to the mentioned literature, this should form a vital assurance against the mistake of mixing in components from the designer's own context, which is the first major cause of design-actuality gaps described in the previous section.

Conceived assumptions about the situation of the user

Still, what is spoken of above mostly concerns the functionality and the user interface of the application. As part of the open source methodology approach I went to Sierra Leone to get rid of possible misconceived assumptions about their situation. I discovered that there are far more fundamental issues that for a long time have been giving the health sector a hard time. Let us compare my findings in Sierra Leone with the following features of my GIS module:

1. The system is web based and can be described as a *thin client*. This means that it relies heavily on its server which may run externally as well as locally. It is light weight and requires minimum of hardware; all you need in order to run it is a web browser. Also, it is operating system independent and can be centrally updated and maintained.
2. Easy to learn and operate. End-users have played a major role in the design process, which should provide for a smooth learning curve for other end-users. As pointed out in the literature (ref 2.2.3), an issue facing developing countries is not so much the access to a particular technology, but dealing with the challenges related to the processes of technological change and the human and social factors that need to be adapted to these processes.
3. Developed for local context. Adapted features like the shapefile generator and the different map source types are crucial elements of the efforts to meet the users' needs in most developing countries.
4. Free and open source software. No expenses involved. Expensive licenses are no longer an issue, and maintenance (ref 2.2.4) can be replicable without incurring large costs as the modification of source code is also free.
5. Integrated in DHIS 2. All GIS data are distributed within the same database as the rest of the system and thus with the same standard for every country. Lippeveld and Sapirie (2000) claim that most developing countries have routine paper-based health information systems in place to collect and report data. These (ref 2.3.1) are seen largely inadequate and ineffective to support health care. The integration of GIS into the DHIS 2 may take advantage of the fact that several developing countries already have invested a lot of effort on entering legacy data

from paper forms into the DHIS 2. Also, Weber (2003) states that the use of standards and open formats instead of data tied to single providers guarantees free access to public information (ref 2.2.4). The implementation of SDMX-HD, the standard format for data exchange in the health domain, in DHIS 2 will soon be finalized, which makes the system inter-operable with other existing applications as well.

I argue that all of the findings from Sierra Leone are handled by the mentioned features of the system:

- *Unavailability of up-to-date shapefiles*. (3 – Local context)

- *Inadequate funding available for training in new complex proprietary solutions, which results in a huge lack of qualified IT personnel. Their slight GIS experiences from the regional offices were bad as they found the current software very hard to learn*. (2 – Easy to learn)

- *Inadequate funding available for acquiring new dedicated high-end hardware, which is required to run most proprietary solutions*. (1 – Web based and light-weight)

- *No IT personnel can be dedicated to install, be trained on the use of, and support new complex proprietary software*. (2 – Easy to learn)

- *Health professionals are over-committed and do not have time for training in GIS software operations with steep learning curves*. (2 – Easy to learn)

- *The GIS currently installed in Mr. Yambashu's computer was proprietary software, which concerned him a lot as they had inadequate funding available for future software licenses and maintenance fees*. (4 – FOSS)

- *The above software could probably have been useful at the SSL, but it was barely used because their computer hardware could hardly deal with such heavy software. They were still practicing paper based map analyses. Performing spatial analysis* *with such computer hardware is very slow*. (1 – Web based and light-weight)

- *Public health users typically make use of only a small fraction of the functionality of large GIS packages*. (3 – Adapted to local context)

- *Public health databases are often in incoherent formats with no standardized geo-referencing, which makes their integration into a GIS difficult*. (5 – Integrated in DHIS 2)

- *The geographical information, sampling data, and the results of field studies are not easily and quickly accessible to public health officials to help them assess and respond to the situation*. (3 – Adapted to local context, 5 – Integrated in DHIS 2)

- *Still, original data are kept in paper records. These have clearly limited distribution and are hard to access. Deterioration and loss are likely to occur*. (1 – Web based and light weight, 5 – Integrated in DHIS 2)

- *Results presented in conferences and journal articles do not provide easy access to the original, unfiltered data to current and future researchers*. (1 – Web based and light weight, 5 – Integrated in DHIS 2)

- *Limited distribution of spatial information and GIS know-how*. (2 – Easy to learn, 5 – Integrated in DHIS 2)

- *There is a need for an efficient, easy-accessible, and user-friendly geospatial information system without steep learning curves, which adequately addresses their specific needs*. (1 – Web based and light weight, 2 – Easy to learn, 3 – Adapted to local context)

- *Agreement on new proprietary GIS solutions would demand significant resources and significant changes to existing plans and IT infrastructures. They expressed huge concern for dependency on proprietary foreign solutions*. (4 – FOSS)

- *Need for technological solutions that do not require large financial, human, and technological resources*. (1 – Web based and light weight, 2 – Easy to learn)

- *Need for technological solutions that are easy to integrate with existing IT infrastructure and capacity*. (1 – Web based and light weight, 4 – FOSS, 5 – Integrated in DHIS 2)

- *Possibility to connect to an external repository that is running a GIS server. This way all GIS users can load shapefiles from this repository into their application, as opposed to their current GIS solution where everyone needs to have the shapefiles on their local hard drive*. (3 – Adapted to local context)

- *Need for data and systems that have the capability of being inter-operable*. (5 – Integrated in DHIS 2.)

It is not reasonable to declare that no proprietary GIS software is easy to learn. Beyond that, however, we can see that FOSS apparently has a lot of advantages over software that is either proprietary or desktop based for this purpose. As stated (ref 2.2.1) by Global Knowledge Partnership (2003), ICT needs to be affordable for the poor, in terms of both acquisition costs and running costs. Weber (2003) says that due to the digital divide (ref 2.2.4) and more specifically the fact that developing countries have limited budgets earmarked for information technology most governments in the developing world are advocating the use of FOSS when it is a feasible alternative to proprietary software solutions. The findings I have mentioned reveal an evident fear of software that is not free of charge as most developing countries hardly can afford high acquisition costs. Gosh et al (2002) make a good point (ref 2.2.4) by stating that whenever the proprietary standards are established, the necessity to follow them is given. Even in an open tender acquisition system, this requirement for compatibility with proprietary standards makes the system biased towards specific software vendors, perpetuating a dependency. Thus, when even the one-off costs are giving developing countries a hard time, I would say that a permanent dependency to high-priced on-going software licenses is obviously a vicious circle.

The second major cause of design-actuality gaps, according to Heeks, is mis-conceived assumptions about the context of the user. I argue that the presence of such assumptions and their influence on my project are, if possible, non-existent, or at least kept to a minimum. I base this on the fact that I through an open source methodology approach have familiarized the actual context by involving end users heavily in the process and being present in a developing country looking for limitations and seen their needs with my own eyes (as described in the empirical part of this paper).

8.2.2 Modifications to suit local needs

According to Sahay and Walsham (1997), an effective implementation of GIS (ref 2.3.2) is severely vulnerable due to the limited availability of useful geographical data. Two aspects lead to this problem; the existence of data and the accessibility of existing data. These aspects are crucial as a GIS application does not make sense without updated data (Saugene 2005). It is reasonable to say that the quality in health care delivery (ref 2.2.2) is largely dependent on the availability of information.

Among my findings in Africa, the unavailability of updated shapefiles is definitely the most crucial issue and a huge problem for any GIS service. This is even more evident in Sierra Leone because of the frequent establishment of new medical health care clinics after the civil war, as described in the Sierra Leonean context chapter. I am not going to discuss the politics in this case, rather the solution I came up with in order to work around it. I realized that this problem would depreciate the value and usability of my project so much that I simply could not accept to settle with it. According to Dr. Pond Sierra Leone was not just an exception. This was actually the situation in most African countries, which motivated me even more.

The result of this is a built-in shapefile generator, which is described in the GIS module chapter. I argue that both the idea and the implementation are major contributions to GIS solutions for developing countries because:

- it eliminates the mentioned issue of not having access to updated shapefiles. I have personally experienced that such files are currently not available for free in Sierra Leone. Now, the maps you want to use in the GIS may be as up to date as you like, as you are able to organize them yourself via the DHIS. The fact that the DHIS is free of charge makes this system available to everyone - not only the highest bidder.

- there is no longer a need to distribute the separate shapefiles/geojson files together with the system. The generator retrieves everything it needs from the DHIS 2 database.
- it improves the thematic maps as you no longer have to deal with ”dead” organisation units. With ”dead” organisation units I mean units that suffer from poor accordance between the shapefile and the database, i.e. missing or duplicated entries, or misspellings. These both disturb the visual presentation and affects the map legend as their value will always be 0. Earlier you had to revise both the database and the shapefile in order to avoid them. Now, the organisation units in the database are the source of both the indicator data and the map, which always gives you 100% accordance.
- the use of the GIS application is made considerably easier as there is no need to register the maps anymore. You are simply asked to select the desired organisation unit level instead of a map containing the organisation units on that level. A map must first be registered correctly as a map object (as described in the GIS module chapter) and then have every organisation unit assigned to the respective organisation unit in the database.
- it heavily speeds up the creation of thematic maps. This is because the calculation algorithm may be a lot simpler and faster as there is no need to match the organisation units in the shapefile against the organisation units in the database.

The obvious disadvantage with this solution is that you need to have the coordinates of the organisation units you want to display stored in your database. This is however a one-off piece of work that everyone can contribute to. Anyone can easily find and report the coordinates of a new medical health care center with a GPS machine. The fact that most of the coordinates of existing medical centers are already available in older shapefiles should make it a manageable task. After that, when a new clinic is opened, it can simply be added as a new organisation unit in the DHIS 2 with its coordinates, and it is automatically ready for use in the GIS, as up to date as it gets. Furthermore, the new clinic would have to be registered in the DHIS 2 anyway, regardless of the GIS, which means that the only extra work needed is the collection of coordinates. Sahay and Walsham (1997) no longer need to worry that (ref 2.3.2) making maps as well as updating them is a costly and time consuming activity and thus difficult in many developing countries due to financial constraints.

A cheap and functional GPS machine used in Mozambique (Saugene 2005).

8.2.3 Open source methodology

To summarize the evaluation section I argue that my project is likely to succeed. I would like to emphasize the fact that open source software development methodology (ref 2.1) has heavily influenced my project in a positive manner. Primarily the principle of *treating users as co-developers* has played a major role. We know that even when effort is made to develop an information system specifically for a developing country, design-actuality gaps may still arise because industrialized-country stakeholders, such as consultants, IT vendors or aid donors, often dominate the IS design process in developing countries (ref 8.1.3). They are likely to inscribe elements from their own context into the design, which usually results in a mismatch between the design and the users' needs. In accordance with the mentioned principle, the participatory design approach (ref 3.2.2) conducted in my project (ref 7.4 and 7.5) has been crucial in the process of counteracting such a mismatch by giving me valuable insight and understanding of the actual context.

The empirical part of this thesis also tells us that another principle of open source methodology, *early releases,* has been carried out. Prototype presentations at early stages have resulted in recruitment of new co-developers (users) that have been included in the development process. As we know, involvement of and collaboration with these co-developers have allowed me to familiarize their context to an extent that would not have been possible sitting in an office in the western world. Furthermore, regular presentations and utilization of the feedback of prototypes have ensured the possibility of creating specific solutions and locally adapted modifications at short notice, as I did not have to stick to a long-term plan. Such solutions and modifications are mentioned (ref 8.1.3) as vital elements of successful information system implementations in developing countries.

The third open source methodology principle mentioned by Robles (2004) is *maximum modularization*. The fact that the DHIS is fully modularized and allows parallel development has provided advantages during the development process as well. The work on the shapefile generator benefited particularly as this solution required extension of the DHIS 2 core in addition to the server part of the GIS module. Parallel work by the core developer team in Oslo actually allowed me to finalize this solution during my stay in Sierra Leone.

Finally, the advantage of releasing software under the BSD licence, which implies that it may be used at no cost, could not be disregarded as a key element in the effort of improving information system implementation sustainability in developing countries.

8.3 Possible limitations

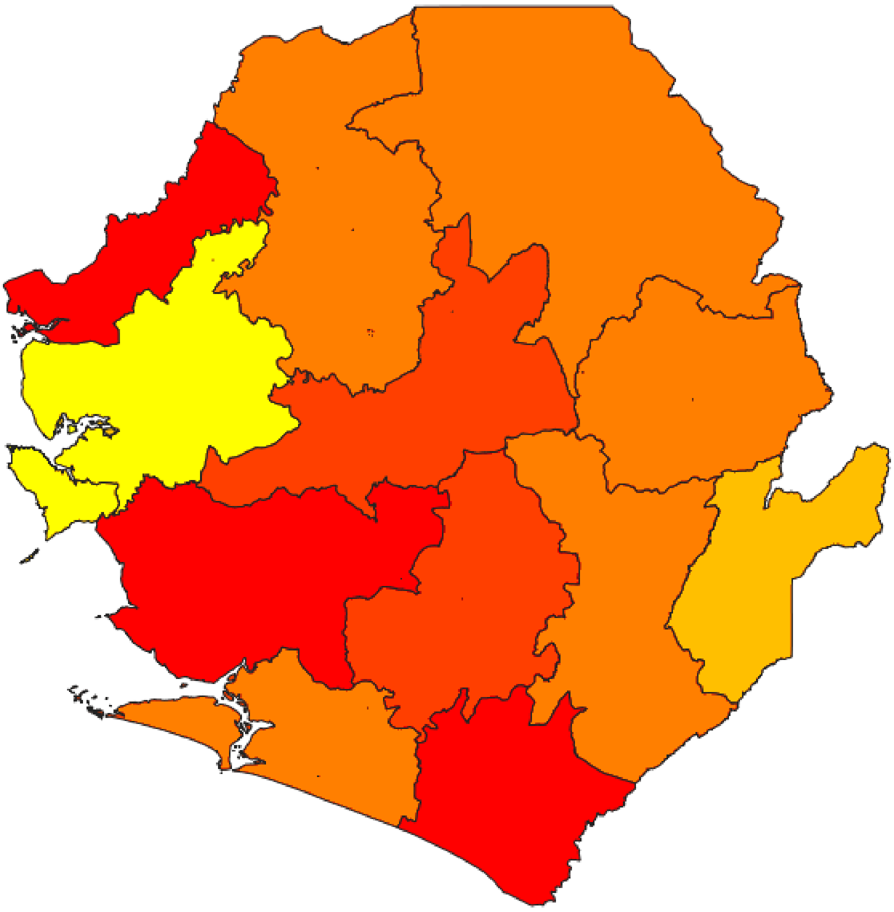
This section focuses on technical limitations related to web based GIS. I will look into the issues that were raised during the workshop in India and decide whether they will cause restrictions in the future.

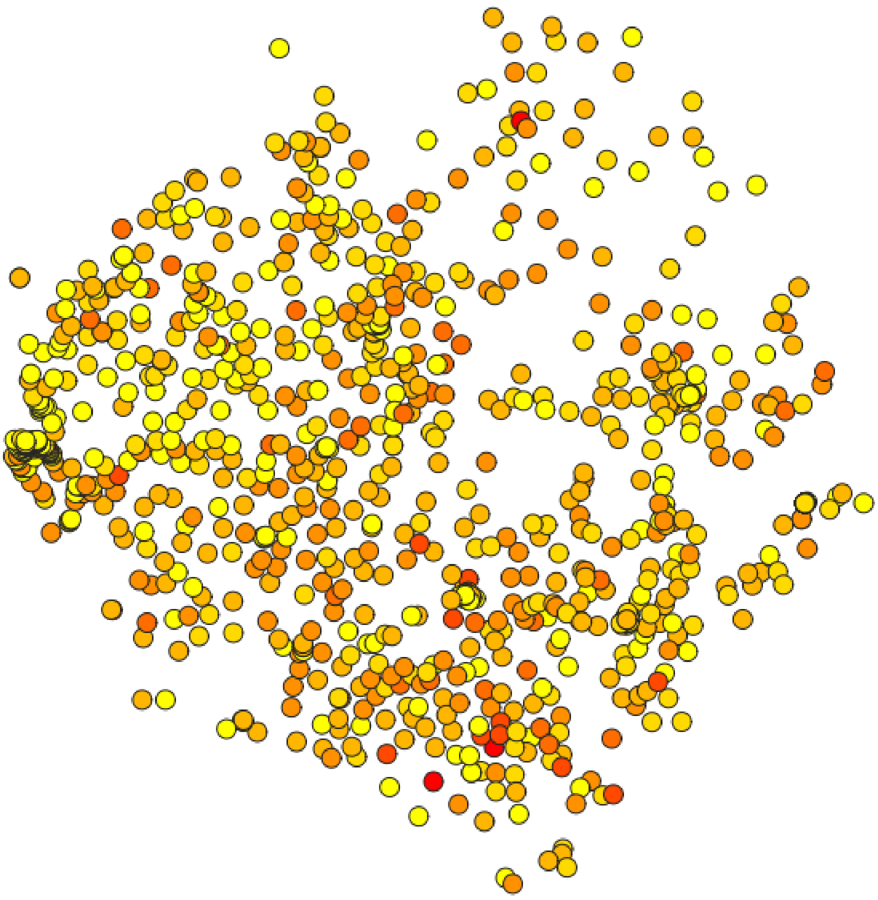
8.3.1 Browser capacity

As discovered during the workshop in India even modern web browsers seem to have trouble handling large amounts of GeoJSON. It would be a rather heavy restriction to all web based GIS applications if it turns out that the browser cannot deal with most of the shapefiles that have been used by desktop GIS software. This is an issue that I have not seen thoroughly discussed anywhere, probably because GeoJSON is still quite new, being finalized in June 2008 (GeoJSON, ref 6.2.2). The web based mapping frameworks that utilize GeoJSON are at a relatively early stage of development as well, which likely explains why I was not able to find any credible literature on the topic. Thus, I find it both interesting in relation to my own project and contributive to the web based GIS movement in general to draw on the experiences I have had over the last year and try to make a conclusion.

Firstly, to be able to benchmark the experiment, I created a test application that reported the number of milliseconds needed to load an un-cached shapefile in GeoJSON format, which is simply JavaScript after all. This seemed to work as planned; the larger GeoJSON file size loaded, the longer time needed. Then, I tried different files to close in on the definite limit of what the browser could handle, which resulted in a quite unexpected outcome; there was not always the bigger file sizes that forced the browser to crash. On other words, there were other factors involved.

To figure this out I contacted developers from Mozilla (Firefox) and Google (Chrome). They both reminded me that OpenLayers uses Scalable Vector Graphics (SVG, ref 6.2.3) to draw the graphics defined in the GeoJSON file, as opposed to e.g. thematic mapping conducted with Keyhole Markup Language (KML), utilized by Google Earth, where proportional symbols are displayed by scaling image icons (Sandvik 2008). This means that the ability of the browser's SVG renderer plays a major role, in addition to the JavaScript engine. Further testing proved that the number of graphical units (called "features" in GeoJSON) heavily influenced the performance in addition to the file size. Have a look at the two following screenshots; districts and health facility clinics in Sierra Leone respectively:





The GeoJSON file size of the districts is about twice as large as the clinics. Still, the clinics take longer to load because of the high amount of features; about 900 versus 13. The large file size of the districts is explained by the rather detailed borders. A border is defined by a series of points (coordinates) that you in the end can draw a line through. A coordinate is represented by two decimal numerals, one for longitude and one for latitude. As opposed to a border (line), a clinic is simply a point. This means that 900 points only require 900 coordinates. If we study the district GeoJSON source we can see that a border, on the other hand, may have tens of thousands of coordinates dependent on the level of detail. This explains the file size difference between the two.

Now, the question is what could be done in order to increase the performance. Removing coordinates from a clinic map is obviously out of the question as it would remove the entire object. This means that we should try to reduce the file size without making the vector data useless. The good news is that files which require heavy SVG rendering (lots of features) do have a small amount of coordinates (points only), while those that are easily SVG rendered (few features, like districts) have lots of coordinates and thus may have their file size reduced considerably. Let us have a look at how this can be done.

Reduce number of decimals

The coordinates from the mentioned shapefiles have nothing less than 15 decimals, e.g. *[-13.274461403113355,8.483453645982836]*. The surface distance per 1 degree change in *latitude* is always approximately 111 km (Wikipedia L 2010), which means they have an accuracy of 0.000000111 mm. The distance per 1 degree change in *longitude* differs from 0 km at the poles (90 and -90 degrees) to 111 km at the equator (0 degrees), which gives at least the same and most likely even better accuracy. This is obviously overkill when we simply want to display polygons that must be recognized as the districts of Sierra Leone and roughly fit the background map. The districts file (first image) has about 20 000 coordinates, i.e. 40 000 longitude and latitude values. By leaving e.g. five decimals and removing the last ten from each value we strip away 400 000 characters, which actually make up about 50% of the file size. And still, our longitudes and latitudes have an accuracy of 1.1 meter. Then, the Pythagorean theorem tells us that the worst case error margin is 1.56 meter, which is more than good enough for the mentioned purpose. A proposed concrete solution could be to create a built-in function that performs such an operation automatically.

Reduce number of coordinates

Instead of editing all the values, another solution might be to make the border less detailed by stripping away e.g. every third or fourth coordinate (only for polygons). This is called shapefile simplification and there are editors that do such operations for us. I have tried a free web application called *MapShaper* (http://mapshaper.com/test/demo.html) that offers three different simplification methods and lets you decide to what extent you want to simplify the borders.

8.3.2 GeoJSON

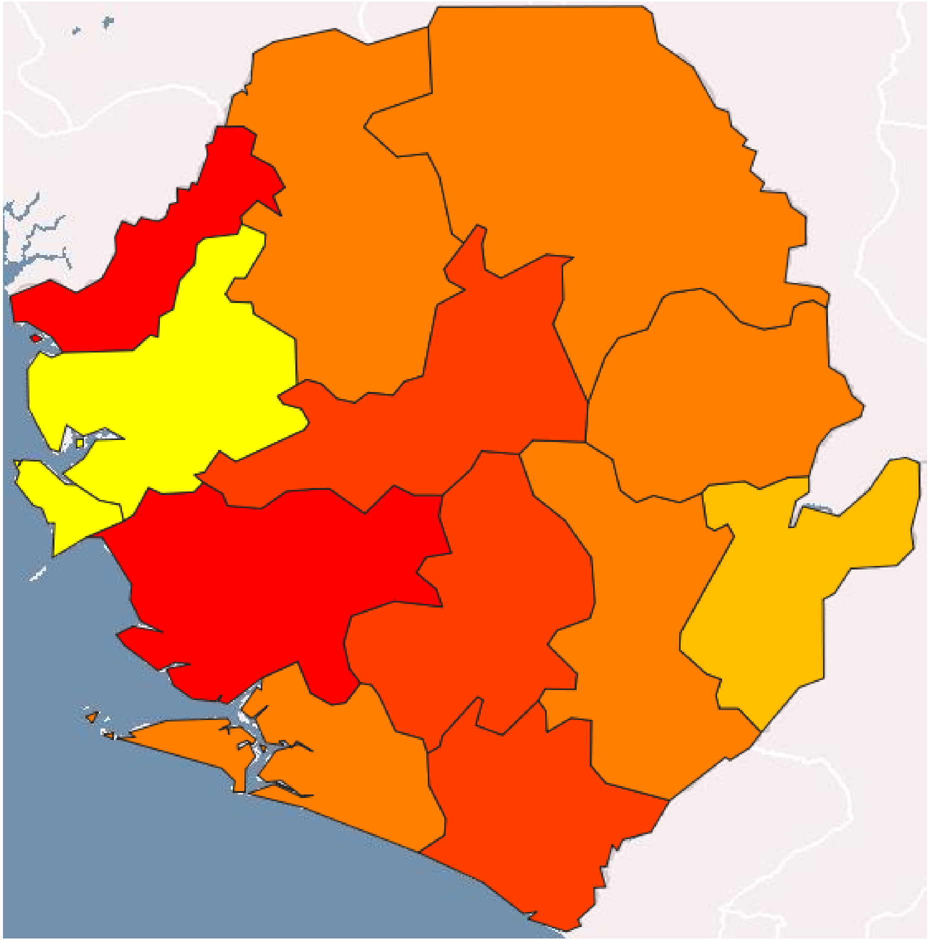
This paragraph has a lot in common with the previous one, but is more related to the GeoJSON format syntax, not only the coordinates. In India a question was raised as to whether GeoJSON produced unnecessarily large files, which is exactly what we want to avoid. Lets have a look at any possible optimization compared to the sample in the GIS module chapter.

The *type* member is required. We want to have several feature objects (e.g. districts and clinics) so we set it to "FeatureCollection". Now, the *features* array is required as well. Any object inside this array has to be of the type feature, so *type* is actually not required. However, there must be an array called *geometry* containing at least a *coordinates* array and a *type* member, because we may have several types of features (e.g. polygons and points) inside our feature collection. The feature object might as well have a *properties* object containing information about the feature (a district in the following sample). We would like to have a property member containing the name of the feature as it is needed both for the user interface and during the creation of thematic maps.

The *crs* object is not required. If our shapefile has a longitude/latitude coordinate system we do not have to define any, as the default system WGS84 uses longitude and latitude as well. Defining a bounding box is not required neither and not needed in most cases. This leaves us with the following valid syntax:

```
{
    "type":"FeatureCollection",
    "features":
    [
        {
            "geometry":
            {
                "type":"Polygon",
                "coordinates":[[[-10,10],[-10,5],[-5,5],[-5,10]]] // dummy square
            },
            "properties":
            {
                "name":"Moyamba"
            }
        }
        // add more feature objects
    ]
}
```

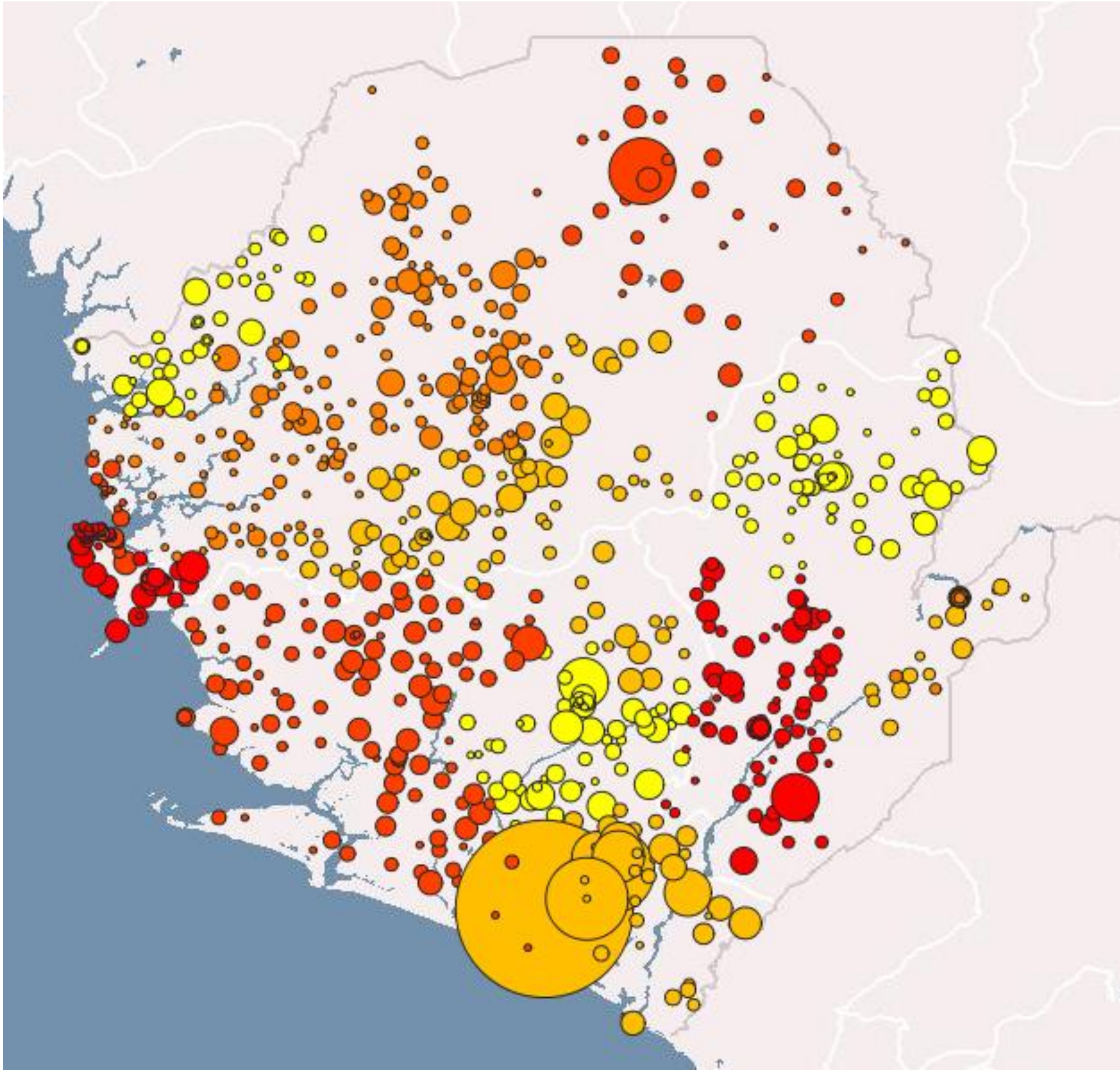
This object contains only the information that is absolutely needed to run my GIS application. In my opinion, the fact that it still validates means that GeoJSON could not be accused of being sub-optimized regarding file size, and thus inefficient. Most JavaScript libraries are even minified before they are published in order to minimize the file size. Minified means that the code style is erased by removing all line breaks and white-spaces. Let us have another look at the map displaying the districts of Sierra Leone after I trimmed its coordinates and decimals, simplified it with MapShaper, minimized the GeoJSON syntax and minified the file:



The file size of this map is 1/40 compared to the first one, which makes it dramatically more responsive and faster to load. Even though the borders are less detailed they still fit the background outlines and are clearly recognizable. With such small file sizes you may apply a whole lot of overlays as well without being afraid of pushing the browsers limits too far. I argue and have shown that with a couple of tweaks *GeoJSON* and even today's *browser capacity* may not cause any restrictions related to web based GIS. Such tweaks can easily be performed with the use of free software like MapShaper and a regular text editor. And the fact that the developers from Mozilla and Google confirmed that improvement of SVG rendering and their JavaScript engine are highly prioritized and under constant development, definitely does not indicate any restrictions in the future.

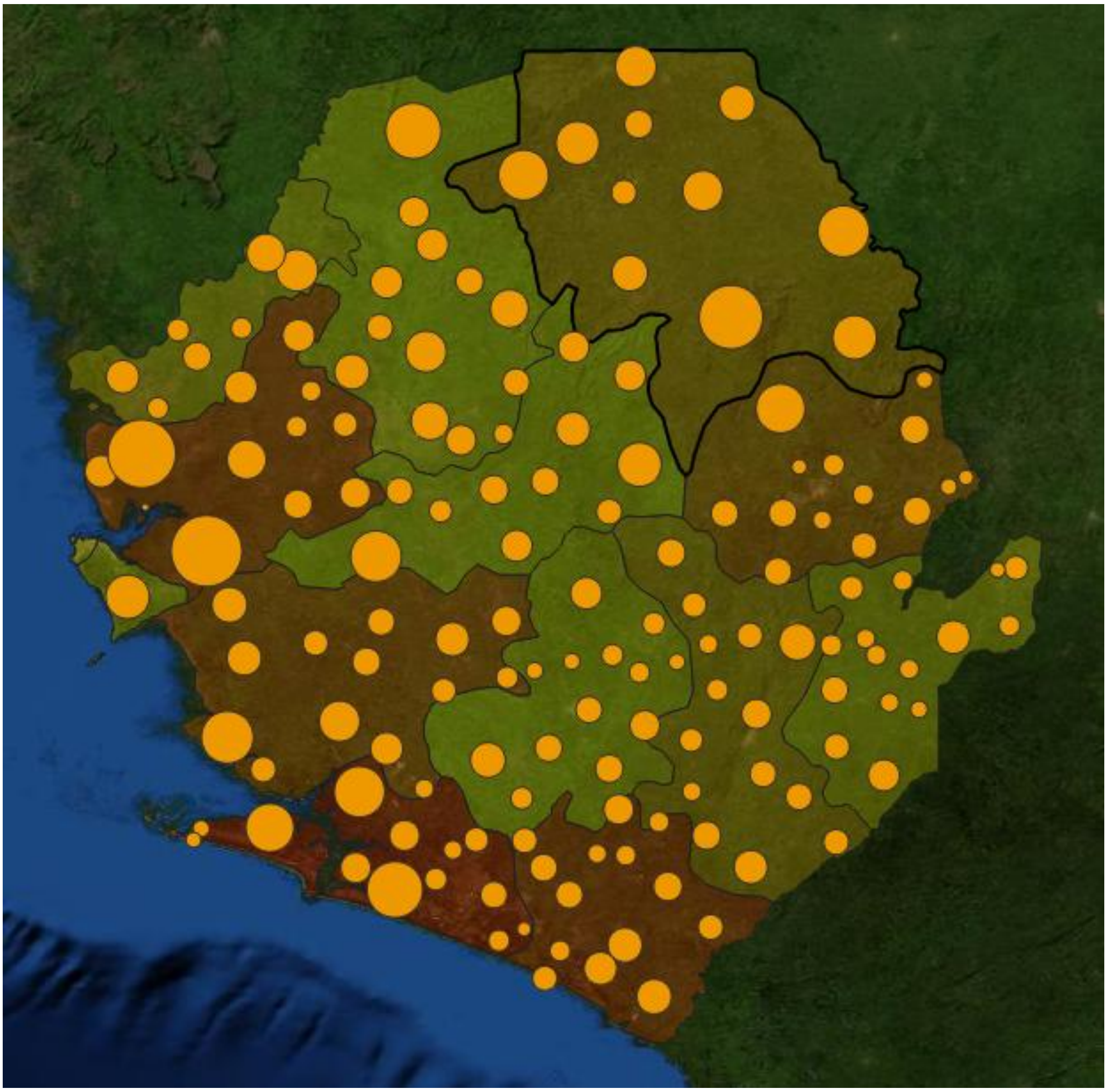
8.4 Thematic mapping

Let us have a look at what our web based GIS application already can do when it comes to thematic mapping, to demonstrate some of the web based GIS capacity that is available. I have currently extended my proportional symbol widget to handle two different indicator values per health facility clinic in order to compare or display more data within the same map:



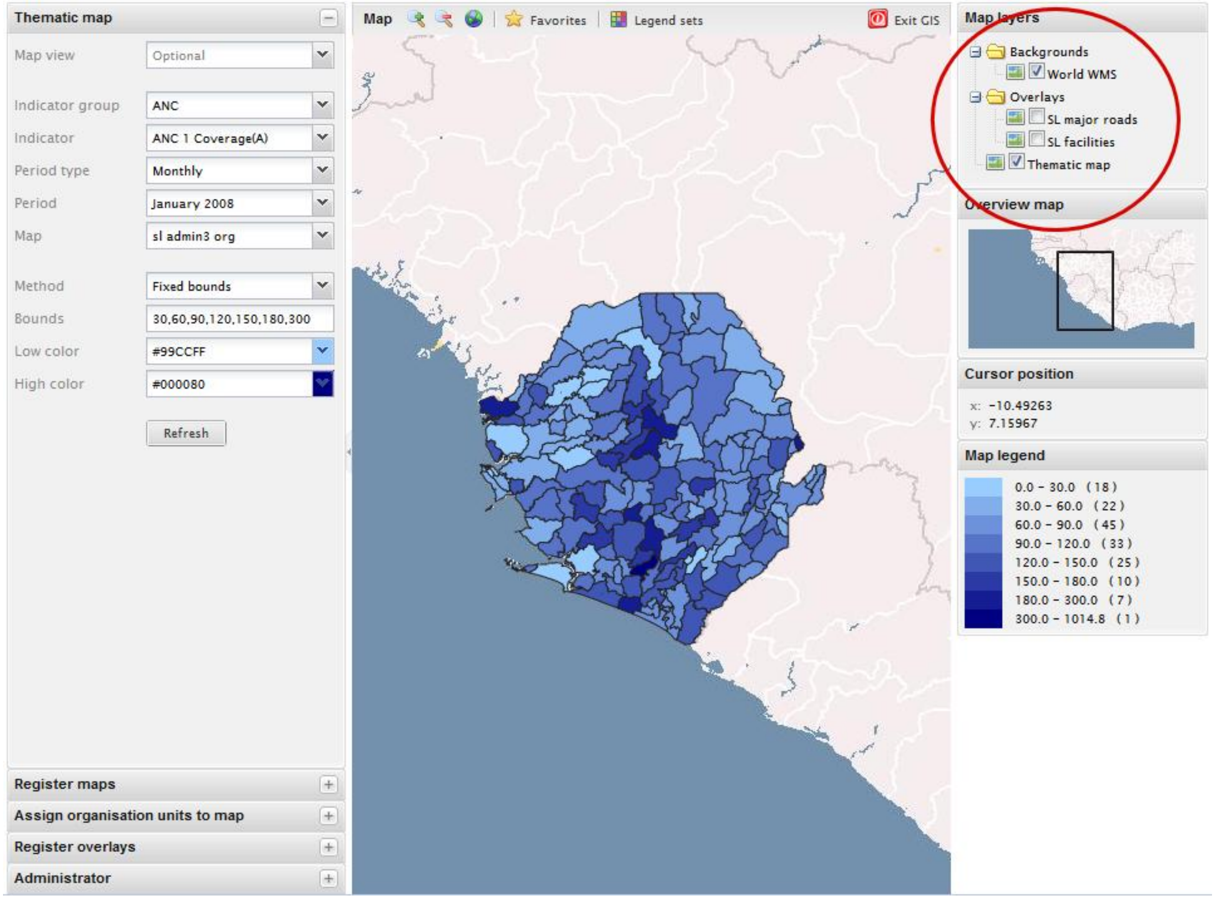
We can see that circles with the same size do not necessarily have the same color. This shows that the generation of size and color are based on different indicators.

We may also display a generated proportional symbol layer on top of a generated polygon layer. By “generated” I mean vector graphics with calculated color or size, which is the opposite of a static layer:

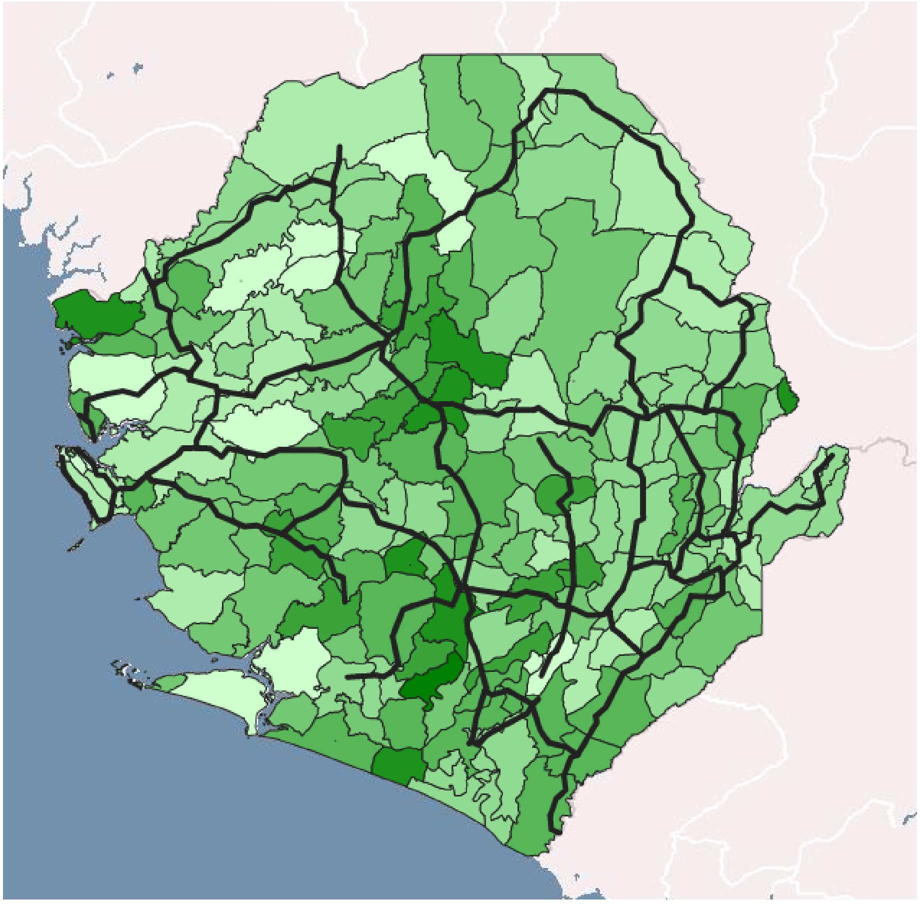


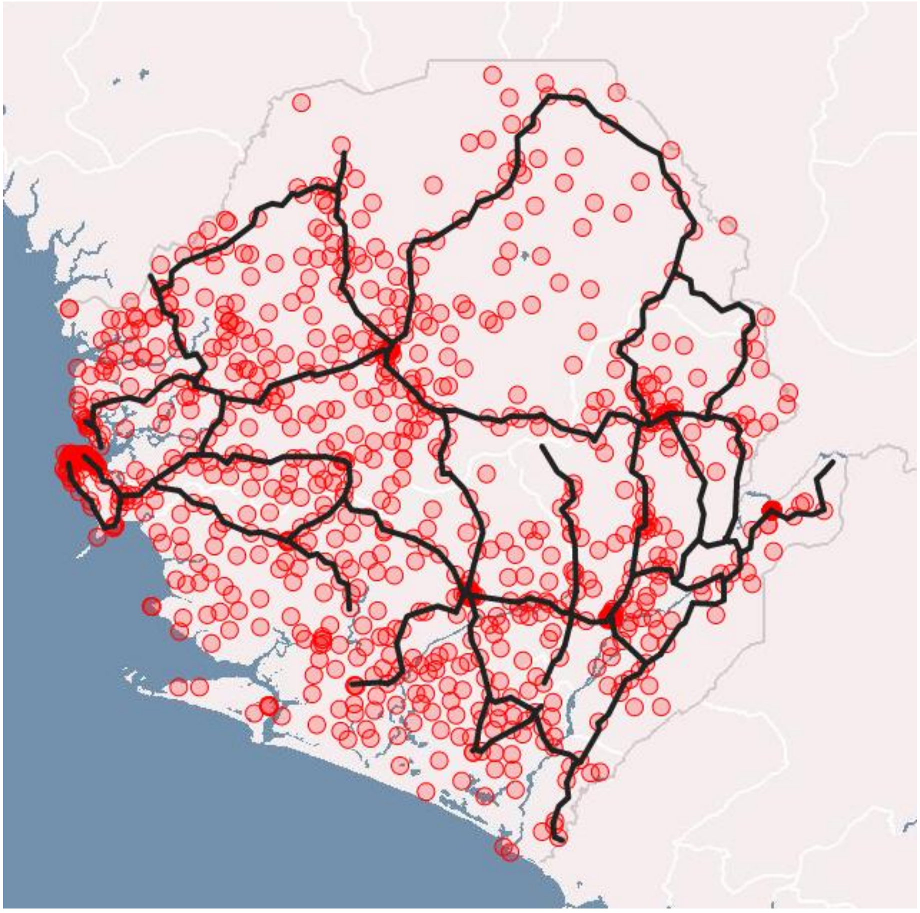
Another way of displaying values could be to adjust the opacity, as demonstrated by the polygons in the image above. The fact that the colors will now bleed into the background, however, is a drawback that generally makes me advise against using this technique.

An Ext layer container will collect all layers added to an OpenLayers map. The new layer container widgets provided by GeoExt make it easy to handle both base layers and overlay layers, as I have implemented in the top right corner:



In other words, we may apply static overlays to the map as well. On the next page, the first image shows the major roads in Sierra Leone as a static layer on top of a generated thematic map. The second image shows static information only, through two different overlays; major roads and health care clinics.





9 Conclusion

In this chapter I summarize my research and discussion and make conclusions to the research questions and finally the research objective.

Research question 1:
Explore the technical capability of web based FOSS GIS development frameworks.

I have explored the first research question at the highest level by using the mentioned development frameworks to build a GIS application myself. I argue and conclude that this technology is already capable of forming a satisfactorily alternative to desktop GIS tools. Instead of simply looking into these frameworks and make assumptions of what they might be capable of, I have personally proven by example that the latest web based GIS tools, which are currently still in development, already have become sufficiently advanced and mature to replicate acknowledged functionality offered by a rich GIS application such as WHO's official desktop GIS tool for developing countries, the HealthMapper. During this process I have received personal training and collaborated with the reputable developers behind the actual mapping frameworks. I believe the insight and understanding this experience has given me makes me entitled to have a qualified opinion in this matter. Additionally, I have contributed to a new and specialized subject area that hardly offers any acknowledged literature or discussion by showing how possible restrictions can be worked around.

Research question 2:
Explore limitations and solutions regarding development and implementation of GIS in developing countries in general, and whether web based FOSS GIS development frameworks and open source methodology can improve sustainability of such implementations.

My second research question has been explored during a one month field study in Sierra Leone. The open source methodology approach has allowed pitfalls related to GIS implementation in developing countries to be detected and avoided. Additionally, my discussion in the previous chapter connects the findings from Sierra Leone and my own

GIS project and reveals the appropriateness of web based FOSS GIS tools in relation to such implementations. This, in addition to the research related to the first question, forms a qualified basis for my conclusion in regard to the main research objective.

My empirical study and discussion identify and emphasize one of the limitations found in Sierra Leone as an issue that in a negative and restrictive way affects health care studies in most developing countries; the unavailability of updated shapefiles. I have contributed both to the technical mindset of GIS and to the health care research in developing countries by developing and implementing a shapefile generator solution, integrated in the DHIS, which deals with this problem. The fact that this was possible by the use of the mentioned tools and frameworks strengthens the conclusion to the first research question.

Research objective:
Explore the capability of web based open source GIS development frameworks and open source methodology to improve sustainability of GIS implementations in developing countries.

The research objective has been addressed by drawing on the experiences acquired during exploration of the research questions. I have shown that the FOSS frameworks are not and will not be a bottleneck in future GIS solutions. On the contrary, I prove that this technology deals with many of the pitfalls related to GIS implementation in developing countries and has in many areas better chances of overcoming them compared to its proprietary, desktop-based alternative.

Furthermore, I have shown that my project is likely to succeed and that the open source methodology approach has provided clear and substantial advantages in the process. The major causes of design-actuality gaps, which are known to be among the main causes for IS failure in developing countries, have been counteracted through adaptation of the software to local context by involving end-users in the design process, high responsiveness to problems and requests for modifications and close follow-up on stakeholders. The project has produced a solution that so far is very well received by a large number of stakeholders and has convinced the WHO to promote it as their official GIS tool for developing countries.

10 Abbreviations

AJAX	Asynchronous JavaScript And XML
API	Application Programming Interface
CSS	Cascading Style Sheets
DHTML	Dynamic HTML
DOM	Domain Object Model
DWR	Direct Web Remoting
EHR	Electronic Health Record
FOSS	Free and Open Source Software
GeoJSON	Geographic JavaScript Object Notation
GDP	Gross Domestic Product
GIS	Geographic Information System
GKP	Global Knowledge Partnership
GML	Geography Markup Language
GSD	Global Software Development
GUI	Graphical User Interface
HIS	Health Information System
HISP	Health Information Systems Programme
HMN	Health Metrics Network
HTML	HyperText Markup Language
II	Information Infrastructure
IS	Information System
IT	Information Technology
JS	JavaScript
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
MDG	Millennium Development Goals
MDX	MultiDimensional eXpressions
NGO	Non-Governmental Organization
OGC	Open Geospatial Consortium
OLAP	Online Analytical Processing
PHC	Primary Health Care
PHU	Peripheral Health Unit

R&D	Research & Development
RIA	Rich Internet Applications
SDC	Swiss Agency for Development and Cooperation
SDMX	Statistical Data and Metadata eXchange
SSL	Statistics Sierra Leone
SQL	Structured Query Language
UNDP	United Nations Development Programme
UNICEF	United Nations Children's Fund
WHO	World Health Organization
WFS	Web Feature Service
WMC	Web Map Context
XML	Extensible Markup Language
WMS	Web Map Service
YUI	Yahoo! User Interface Library

11 References

Aday, L., Anderson, R., 1981. Equity of access to medical care: a conceptual and empirical overview. Med. Care 19(Suppl.): 4–27.

Akrich, M., 1992. The de-scription of technical objects. In Shaping technology/building society, eds.W. E. Bijker and J.Law.Cambridge, MA: MIT Press.

Avgerou, C., and Walsham, G., 2000. Information technology in context. Aldershot, UK: Ashgate.

Avison, D., & Fitzgerald, G., 2003. Information Systems Development Methodologies, Techniques and Tools (3rd Edition ed.). London: McGraw-Hill.

Avison D., Lau F., Myers M., and Nielsen P. A., 1999. "Action Research" in Communications of the ACM, Vol. 42.

Baark, E., and Heeks, R., 1999. Donor-funded information technology transfer projects. Information Technology for Development 8(4):185–197.

Bansler, J., 1987. "Systemutvikling, teori og historie i skandinavisk perspektiv," *Studentlitteratur*, Lund.

Baskerville, Richard L., 1999. "Investigating Information Systems With Action Research": Communications of the Association for Information Systems, Vol. 2, Article 19.

Battin, R. D., Crocker, R., Kriedler J., and Subramanian K., 2001. "Leveraging Resources in Global Software Development," *IEEE Software*, pp. 70-77.

Bijker, W. E., and Law, J., 1992. General introduction. In Shaping technology/building society, eds. W. E. Bijker and J. Law, pp. 1–14. Cambridge, MA: MIT Press.

Bjørn-Andersen, N., & Hedberg, C., 1997. Designing Information Systems in an Organizational Perspective.

Boehm, B., 1981. Software engineering economics. New York: Prentice Hall.

Boland R., and Fitzgerald, B., 2004. "Transitioning from a Co-Located to a Globally-Distributed Software Development Team: A case study at analog devices inc."

Braa, J., 1997. Use and Design of Information Technology in Third World Contexts with a focus on the Health Sector, Case Studies from Mongolia and South Africa, Thesis for PhD. University of Oslo.

Braa, J., and Hedberg, C., 2002. Developing district-based health care information systems. The Information Society 18(2):113–127.

Braa, J., Monteiro, E., and Reinert E., 1995. ”Technology transfer vs. technological learning: Itinfrastructure and health care in developing countries,” *IT for Development*, pp. 15-23.

Braa, J., Monteiro, E., and Sahay S., 2004. “Networks of Action: Sustainable health information systems across developing countries”, *MIS Quarterly*, pp. 337-362.

Brooks, F., 1995. The Mythical Man-Month: Essays on Software Engineering, Anniversary Edition (2nd Edition).

Budic, Z., and David, G., 1994. Implementation and Management effectiveness in adoption of GIS technology in Local Governments.

Burrough, P., 1986. A Principles of Geographical Information Systems for Land Resources Assessment, Clarendon Press, Oxford.

Ciborra, C. U., 2000. “From control to drift. The dynamics of corporate information infrastructures”, *Oxford: Oxford University Press*.

Clarke, K.C., McLafferty, S.L., Tempalski, B.J., 1996. On epidemiology and geographic information systems: a review and discussion of future directions, in Emerging Infectious Diseases 2, 85±92.

Clemons, E. K., Thatcher, M. E., and Row, M. C., 1995. Identifying sources of reengineering failures: A study of the behavioral factors contributing to reengineering risks. Journal of Management Information Systems 12(2): 9–36.

Cornford T., and Smithson S., 1996. “Project research in information systems – A student’s guide,” MacMillan Press.

Croner, C.M., 2003. Public Health, GIS, and the Internet, National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Annu. Rev. Public Health 2003. 24:57–82 doi: 0.1146/annurev.publhealth.24.012902.140835

Damian D., and Zowghi D., 2002. "The impact of stakeholders' geographical distribution on managing requirements in a multisite organization," in *Proceedings IEEE Joint International Conference on Software Engineering*, Los Alamitos.

The Economist, 2000. No gain without pain, Government and the Internet Survey. 24 June: 7-10.

Ehn, P., 1993. Scandinavian design: On participation and skill. In: Schuler, D. & Namioka, A. (eds) Participatory Design. Principles and Practices.

Georgiadou, Y., Puri, S., Sahay, S., 2005. Towards a Research Agenda to Guide the Implementation of Spatial Data Infrastructures – A Case Study from India.

Gosh R., Krieger B., Glott R., and Robles G., 2002. "Free/Libre and Open Source Software: Survey and Study," *International Institute of Infonomics*, University of Maastricht, Netherlands.

Greenbaum, J., & Kyng, M., 1991. Design at Work. Cooperative design of computer systems.

Greenwood, D., and Levin M., 1998. "Introduction to Action Research: Social Research for social change," SAGE publications.

Hanseth, O., and Monteiro, E., 1997. "Inscribing behaviour in information infrastructure,"*Accounting, Management and Information Technologies*, pp. 183-211.

Heeks, R., 1999. Information and Communication Technologies, Poverty and Development. Development Informatics Working Paper Series, (5).

Heeks, R., 2000. Information Systems and Developing Countries: Failure, Success, and Local Improvisations.

Herbsleb, J. D., and Grinter, R. E., 1999. "Splitting the organisation and integrating the code: Conway's Law revisited," in *Proceedings from the 21st International conference on software engineering*, New York: ACM Press, pp. 85-95.

Herbsleb, J. D., Mockus, A., 2003. "An empirical study of speed and communication in a globallydistributed software development," *IEEE Transactions on Software Engineering*.

Høyer R., 1971. "Databehandling, organisasjon og bedriftsmiljø," *Norsk Produktivitetsinstitutt i Oslo*.

James, G., 1997. IT fiascoes . . . and how to avoid them. Datamation. November: 84–88.

Kenaroglu, B., 2000. Implications of Information Technology in Developing Countries and Its Impact in Organizational Change. Turkish Aerospace Industry.

Kitiyadisai, K., 2000. The implementation of IT in reengineering the Thai Revenue Department. In Information flows, local improvisations and work practices. Proceedings of the IFIP WG9.4 Conference 2000. Cape Town: IFIP.

Kling R., Scacchi W., 1980. "The web of computing: computer technology as social organisation" in *Advances of Computers*, pp. 61-110.

Korac-Boisvert, N., and Kouzmin, A., 1995. Transcending soft-core IT disasters in public sector organizations. Information Infrastructure and Policy 4(2):131–61.

Law, J., 1986. "On the methods of long-distance control: Vessels navigation and the Portuguese route to India".

Longley, P., Goodchild, M., Maguire, D., & Rhind, D., 2001. Geographic Information Systems and Science: John Wiley & Sons, LTD.

Lovett, A., Haynes, R., Bentham, G., Gale, S., Brainard, J., Suennenberg, G., 1998. Improving health needs assessment using patient register information in a GIS.

Lyytinen, K., and Hirschheim, R., 1987. Information systems failures: A survey and classi cation of the empirical literature. Oxford Surveys in Information Technology.

Malone, T., & Crowston, K., 1991. Toward an Interdisciplinary Theory of Coordination. MIT: Center for Coordination Science.

Mangesi, K., 2007. ICT in Education in Sierra Leone, Survey of ICT and education in Africa: Sierra Leone Country Report.

Macome, E., 2003. The dynamics of the adoption and use of ICT-based initiatives for development: results of a field study in Mozambique. Unpublished Ph.D., University of Pretoria, Pretoria.

McLafferty, S.L., 2003. GIS And Health Care, Annu. Rev. Public. Health. 2003.24:25-42.

Mohan, J., 1993. Healthy indicators? Applications of census data in health care planning. In Population Matters: The Local Dimension, ed. A Champion, pp. 136–49. London: Paul Chapman.

Mosse, E., & Sahay, S., 2003. Counter Networks, Communication and Health Information Systems: A Case Study from Mozambique.

Moussa, A., and Schware, R., 1992. Informatics in Africa. World Development 20(12): 1737–1752.

NORAD, 2002. ICT in development co-operation.

Nordal, K., 2006. The Challenge of Being Open - Building an Open Source Development Network.

Odedra-Straub, M., ed. 1996. Global information technology and socioeconomic development. Nashua, NH: Ivy League.

Peters, G., and Hall, G.B., 1999. Assessment of ambulance response performance using a geographic information system. Soc. Sci. Med. 49:1551–66

RHINO, 2001. Routine Health Information Network: The Potomac Statement on Investment in Routine Health Information in Developing Countries.

Robles, G., 2004. "A Software Engineering Approach to Libre Software".

Roche, E., and Blaine, M., eds. 1996. Information technology, development and policy. Aldershot, UK: Avebury.

Rolland K., and Monteiro E., 2002. "Balancing the local and the global in infrastructural information systems," *The Information Society*.

Sahay S., 1998. "Implementing GIS technology in India: Some issues of time and space," *Accounting, Management, and Information Technology*, pp. 147-188.

Sahay, S., Madon, S., 1997. Managing natural resources using GIS: Experiences in India, Information and Management, Elsevier Science, 0378-7206/97.

Sandvik, B., 2008. Using KML for Thematic Mapping.

Sauer, C., 1993. Why information systems fail: A case study approach. Henley, UK: Alfred Waller.

Sauer, C., 1999. Deciding the future for IS failures: Not the choice you might think. In Rethinking management information systems, eds. R. Galliers and W. L. Currie, pp. 279–309. Oxford, UK: Oxford University Press.

Sauerborn, R., Karam, M., 2000. Geographic Information System, in Design and Implementation of health information system.

Saugene, Z., 2005. Challenges, Opportunities and Strategies for Using Geographic Information Systems for Public Health Management. An Action Research Study from Mozambique.

Snow, R.W., Gouws, E., Omumbo, J., 1998. Models to predict the intensity of Plasmodium falciparum transmission: applications to the burden of disease in Kenya, Transactions of the Royal Society of Tropical Medicine and Hygiene 92, 601±606.

Suchman, L., 1987. Plans and situated actions. Cambridge, UK: Cambridge University Press.

Shields, P., and Servaes, J., 1989. The impact of the transfer of information technology on development. The Information Society 6(1–2): 47–57.

Tanser, F., Wilkinson, D., 1999. Spatial implications of the tuberculosis DOTS strategy in rural South Africa: a novel application of geographical information system and global positioning system technologies, Trop. Med. Int. Health 4:634– 38.

Taylor, D.R.F., 1991. GIS in Developing Nations- Special issue, International Journal of Geographical Information Systems, 5(1).

Teeffelen, P., Leo, G., & Otto, V., 1993. Possibilities and Constraints of GIS applications in developing countries.

Wadsworth, Y., 1998. What is Participatory Action Research? Action Research International, Paper 2.

Walsham, G., 2001. Making a World of Difference: IT in a Global Context, John Wiley & Sons Ltd ISBN 0-47187724-7.

Wang, L., 2007. Education in Sierra Leone: Present Challenges, Future Opportunities, World Bank Publications, p. 1,2,3.

Weber S., 2003. "Open Source Software in Developing Economies", *University of California*, Berkeley.

Weill P., and Broadbent M., 1998. "Leveraging the new infrastructure – How market leaders capitalize on information technology," *Boston Harvard School Press*.

Wynn, E., and deLyra, J., 2000. A strange attractor in the chaos of global IS development. In Information flows, local improvisations and work practices. Proceedings of the IFIP WG9.4 Conference 2000. Cape Town.

Yates, J., 1989. "Control through communication: The rise of system in America", *Baltimore, MD: Johns Hopkins University Press*.

Yeh, A.G., 1991. The development and applications of geographic information systems in developing countries, International Journal of Geographical Information Systems, 5(1), pp.5-27.

Øverland, L., 2006. Global Software Development and Local Capacity Building: A means for improving Sustainability in Information Systems Implementations.

11.1 Internet references

Bazaar 2010
http://bazaar-vcs.org/en/
Bazaar official website

Camptocamp 2010
http://www.camptocamp.com/
Camptocamp official website

Canonical 2010
http://www.ubuntu.com/news/canonical-open-sources-launchpad
"Canonical releases source code for Launchpad", Ubuntu official website

CIA 2010
https://www.cia.gov/library/publications/the-world-factbook/fields/2004.html
CIA World Factbook

Direct Web Remoting 2010
http://directwebremoting.org
Direct Web Remoting official website

Examiner 2009
http://www.examiner.com/x-19273-San-Jose-Scholarly-Research-Examiner~y2009m9d24-Interview-types-Structured-semistructured-and-unstructured
Interview types

GeoExt 2010
http://www.geoext.org/
GeoExt official website

GeoJSON 2010
http://geojson.org/
Official GeoJSON website

GKP 2003
http://www.globalknowledge.org/isct4d/index.cfm
ICT4D - Connecting people for a better world

H2 2010
http://www.h2database.com/html/main.html

H2 Database Engine official website

HealthMapper 2010
http://www.who.int/health_mapping/tools/healthmapper/en/index.html
Official HealthMapper website

Jira 2006
http://www.atlassian.com/software/jira/
Atlassian Jira official website

JSON 2010
http://json.org/
Official JSON website

Kambia 2010
http://www.kambia.org.uk/information_about/sierra_leone.htm
”The Kambia Appeal – Improving health in Sierra Leone”

Launchpad 2010
https://launchpad.net/
Launchpad official website

Mackenzie 2007
http://news.bbc.co.uk/2/hi/programmes/newsnight/6231905.stm
Mackenzie J., ”Sierra Leone's failing health”, BBC

MapFish 2010
http://www.mapfish.org/
MapFish official website

Maven 2006
http://maven.apache.org/
Apache Maven project official website

OpenLayers 2010
http://openlayers.org/
OpenLayers official website

Raymond 2010
http://catb.org/~esr/writings/cathedral-bazaar/cathedral-bazaar/index.html
The Cathedral and the Bazaar

Schuler 2008
http://www.publicsphereproject.org/patterns/print-pattern.php?begin=36
Participatory design

SL Encyclopedia 2008
http://www.daco-sl.org/encyclopedia/1_gov/1_7njala.htm
"Njala University College", Sierra Leone Encyclopedia

Spring 2006
http://www.springframework.org/
Spring Framework official website

Statistics Sierra Leone 2010
http://www.statistics.sl/
Official Statistics Sierra Leone website

Struts 2010
http://struts.apache.org/
Official Struts website

SVG 2010
http://www.w3.org/Graphics/SVG/
SVG official website

UN 2000
http://www.un.org/milleniumgoals/
The Millennium Goals

UNDP 2005
http://sdnhq.undp.org/it4dev/
Information and Communication Technology for development

UNDP 2009
http://hdrstats.undp.org/en/indicators/20.html
"Human Development Report 2009", United Nations Development Programme

U.S. Department of Labour 2002
http://www.dol.gov/ilab/media/reports/iclp/tda2001/Sierra-leone.htm
" 2001 Findings on the Worst Forms of Child Labor"

Velocity 2006
http://jakarta.apache.org/velocity/
Apache Jakarta Velocity project official website

Walsh 2008
http://news.bbc.co.uk/2/hi/health/7202278.stm
Walsh, F., "Survival is tough in Sierra Leone", BBC

Wikipedia A 2010
http://en.wikipedia.org/wiki/Sierra_leone
Sierra Leone

Wikipedia B 2010
http://en.wikipedia.org/wiki/Healthcare_in_Sierra_Leone
Healthcare in Sierra Leone
Wikipedia C 2010
http://en.wikipedia.org/wiki/Launchpad_%28website%29
Launchpad

Wikipedia D 2010
http://en.wikipedia.org/wiki/Junit
Junit

Wikipedia E 2010
http://en.wikipedia.org/wiki/Ext_JS
Ext

Wikipedia F 2010
http://en.wikipedia.org/wiki/OLAP_cube
OLAP cube

Wikipedia G 2010
http://en.wikipedia.org/wiki/Jetty_(web_server)
Jetty (web server)

Wikipedia H 2010
http://en.wikipedia.org/wiki/Json
Json

Wikipedia I 2010
http://en.wikipedia.org/wiki/World_Health_Organization
World Health Organization
Wikipedia J 2010
http://en.wikipedia.org/wiki/HMN
HMN

Wikipedia K 2010
http://en.wikipedia.org/wiki/Participatory_design
Participatory design

Wikipedia L 2010
http://en.wikipedia.org/wiki/Latitude
Latitude

Wikipedia M 2010
http://en.wikipedia.org/wiki/Open_source
Open source

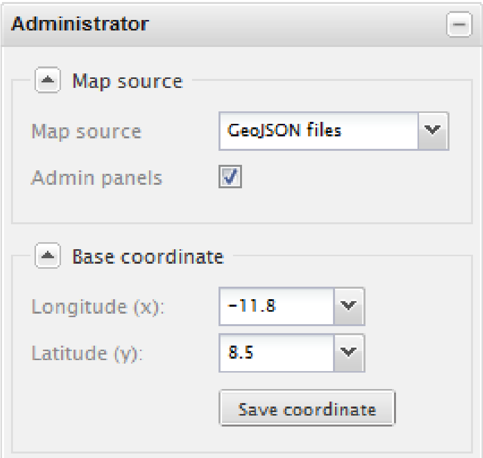
Wikipedia N 2010
http://en.wikipedia.org/wiki/Free_software
Free software

World Refugee Survey 2009
http://www.refugees.org/article.aspx?id=2324
Refugees

Appendix A

This appendix presents the GIS user documentation. A demo of the GIS is currently running at http://demo.dhis2.org and may be accessed with username *admin* and password *district*.

Administrator panel



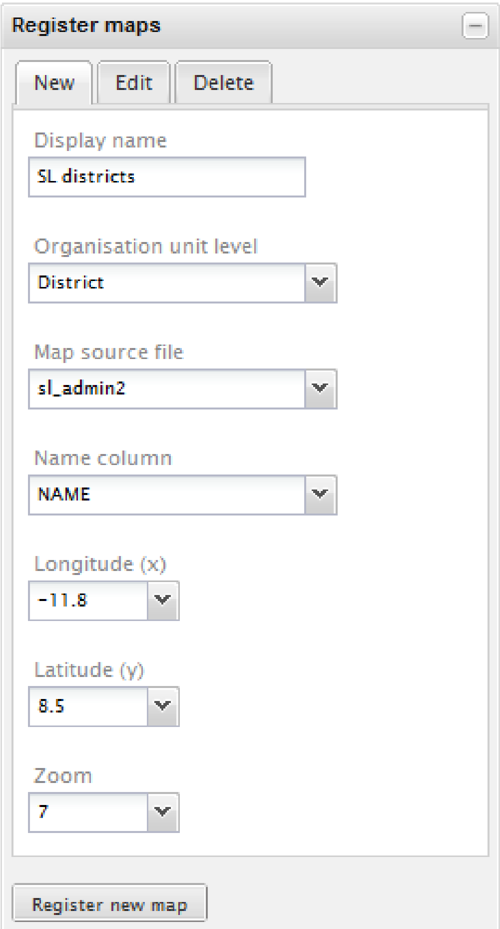
- *Map source*
    - *GeoJSON files*
      If the map source is set to *GeoJSON files* you will find the maps registered from local GeoJSON files in the *Map* combo box in the *Thematic map* panel. The Admin panels check box will become visibile.
    - *Shapefile*
      Maps registered from a locally or externally running Geoserver will appear in the *Map* combo box in the *Thematic map* panel. The Admin panels check box will become visibile.
    - *DHIS Database*
      The *Map* combo box will simply be populated by the existing organisation unit levels and shapefiles will be created by the application on the fly. Organisation units must have coordinates stored in the datasbase in order to be displayed in the map. This function is mainly intended for the facility level as it is easy to maintain and thus will offer up-to-date shapefiles.

- *Admin panels*
  Show/hide the map and overlay management panels; *Register maps, Assign organisation units to map* and *Register overlays*.

- *Longitude (x) / Latitude (y):*
  The base coordinates for the specific country will appear as default when you register a new map. Place the cursor in the visually estimated center of the country and note the coordinates displayed in the *Cursor position* panel to the right:

Register maps

Create a map by registering a GeoJSON file:



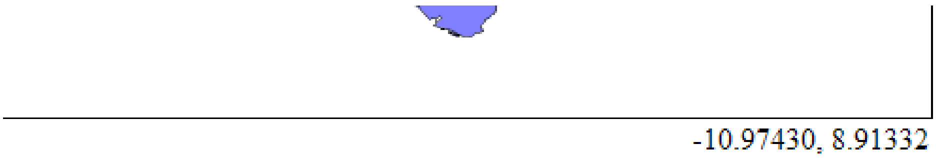
- *Display name*
  This name represents your map in the *Map* combo box in the *Thematic map* panel.

- *Organisation unit level*
  The level of the organisation units displayed in the map.

- *Map source file*
  A list of geojson files placed in the reference folder (DHIS2_HOME/geojson/) will appear in the combo box.

- *Name column*
  A list of all column names in the selected geojson file will appear in the combo box. The column you select will be matched against DHIS organisation unit names. There are several ways to get the right one, e.g.

  - with Geoserver:



ID	ADM1_NAME	NAME	ISO_CTRY	POPDENS90	TOT_POP90	ADM1CODE	LVLID
	Northern	Koinadugu	SL	15	190	69403	SLP001002000000000

  - by opening the .dbf file:

	B	C	D	E	F	G	H
ID	ADM1_NAME	NAME	ISO_CTRY	POPDEI	TOT_POI	ADM1	LVLID
101	Southern	Bo	SL	37	213	69401	SLP00200400(
102	Southern	Bonthe	SL	25	89	69401	SLP00200200(
103	Southern	Moyamba	SL	29	202	69401	SLP00200300(

  - directly from the GeoJSON file:

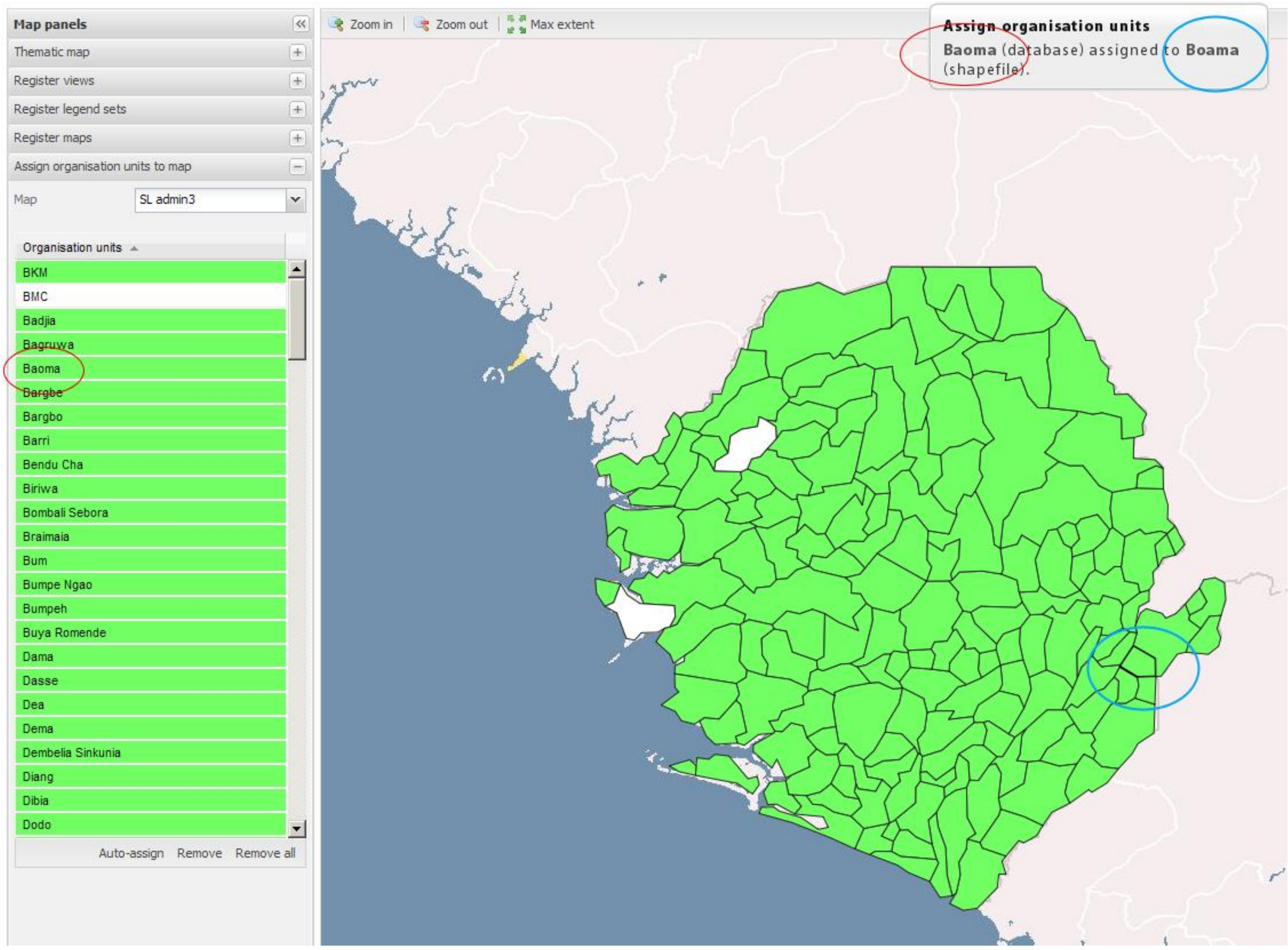

```
084505753,9.33470018916898],[-10.663308431745355,9.300
.612967389934997,9.066730230109561],[-10.58479703041165
"Northern","NAME":"Koinadugu","ISO_CTRY":"SL","POPDENS
admin2.11","geometry":{"type":"MultiPolygon","coordinat
38628116702,8.563647919354837],[-13.065101097857601,8.
```

- *Longitude / latitude*
  The *base coordinates* will appear as default, but you are free to enter new ones.

- *Zoom*
  Enter the default zoom level. 7 is usually a good choice for countries.

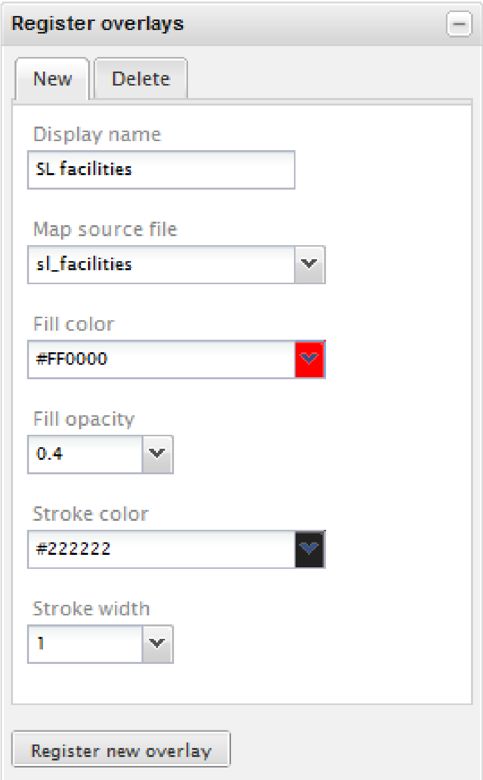
Assign organisation units to map

Select a registered map and wait for it to load. The organisation units (OU) in your database on this level will appear in the list and colors will appear in the map. What we want to do here is creating relations between OUs in the database and the corresponding OUs in the shapefile. First, try *Auto-assign* at the toolbar below the list of OUs to let the application link the OUs with a matching OU name in the shapefile for you. The polygons that remain white you will have to link manually by first selecting a white OU in the list and then click the corresponding OU in the map.



The *remove* button at the botton toolbar removes the link of the OU you have selected in the list and the *remove all* button removes all OU links for the selected map.

Register overlays

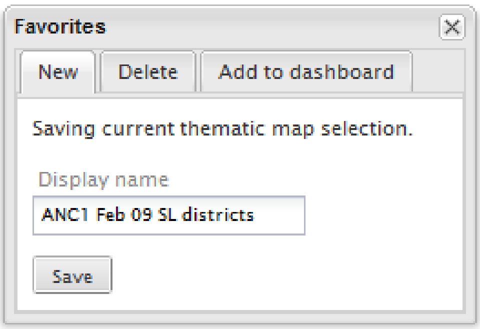


- *Display name*
  Represents your overlay in the layer tree in the upper right corner.
- *Map source file*
  The GeoJSON file name.
- *Fill color*
  Decides the fill color if the overlay has polygons or points.
- *Fill opacity*
  Select an opacity level between 0 (invicible) and 1 (solid).
- *Stroke color*
  The stroke color over lines and polygon borders.
- *Stroke width*
  Select a stroke width.

Thematic map

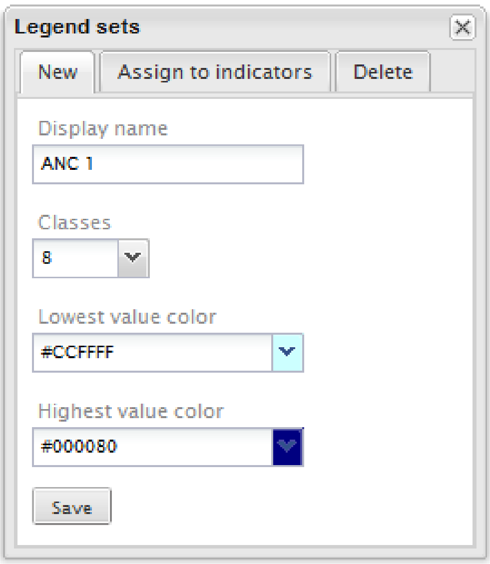
This panel should be rather self-explanatory . Calculation *method* alludes to the legend interval size and set to *Equal intervals* they will be “highest map value – lowest map value / number of classes”. Choose *Fixed bounds* and you may set your own legend limits, e.g. “20,40,60”. The *classes* box decides the number of intervals.

Register views



Save the current thematic map view in order to restore it whenever you want via the *Map view* combo box in the *Thematic map* panel. By adding your views to *DHIS 2 Dashboard* you may access them directly from the dashboard by inserting *Map views* into one of link areas.

Register legend sets



A legend set may be connected to many indicators, but an indicator may only have one legend set. Thus, you may select many indicators when you create a legend set. When an indicator with an assigned legend set is selected in the *thematic map* panel, the number of classes, low color and high color is automatically applied. To assign a legend set to one or many indicators click the *Assign to indicators* tab, select one or many indicators in the list (use the *Shift* button to select a group and the *Ctrl* button to pick one by one) and finally click the *Assign to indicators* button.