Healthsites.Io - A Free, Curated, Global, Canonical Source of Healthcare Location Data for Emergency Relief, Disease Epidemic and Crisis Situations

Dražen ODOBAŠIĆ[1] and Franka GRUBIŠIĆ[2], Croatia

Keywords: healthcare location data, emergency relief, disease epidemic, crisis situations, online maps, easily accessible locations

SUMMARY

When disaster strikes, the interaction between endangerment and disaster may result in significant injuries and possibly, the loss of human lives. In all emergency situations, the prioritisation of health services must focus on meeting both the short-term and long-term needs of the victims. Countries are being encouraged to improve their disaster preparedness, along with growing international commitment to strengthening health systems.

Currently when a natural disaster or disease outbreak occurs there is a rush to establish accurate health care location data that can be used to support people on the ground. This has been demonstrated by events such as the Haiti earthquake and the Ebola epidemic in West Africa. As a result valuable time is wasted establishing accurate and accessible data.
healthsites.io establishes this data and the tools necessary to upload, manage and make the data easily accessible.

Currently health care location data is either incomplete or not available in an accessible open manner. Healthsites.io aim to provide a single point of reference for healthcare workers, aid agencies, contingency planners, government agencies and citizens who need access to a highly curated global and community validated dataset of healthcare facilities. By simply clicking on the map users can discover what healthcare facilities exist at any global location and find out what resources exist at any individual healthcare facility.

[1] Faculty of Geodesy University of Zagreb, e-mail: dodobas@geof.hr
[2] Faculty of Geodesy University of Zagreb, e-mail: fgrubisic@geof.hr

Healthsites.Io - A Free, Curated, Global, Canonical Source of Healthcare Location Data for Emergency Relief, Disease Epidemic and Crisis Situations

Dražen ODOBAŠIĆ[3] and Franka GRUBIŠIĆ[4], Croatia

1. INTRODUCTION

Crisis situations create serious problems all around the world. They are inherently linked with land and thus the surveying profession. Over the last couple of years Volunteered Geographic Information (VGI) have emerged as an important additional source of information in disaster management (Dijkstra, Grubišić, Leitz, Kagerah, & Unger, 2015). Geographic data has become more opened, and with the rise of crowdsourcing mapping activities, geographic data has become more accessible for humanitarian actors. This is in line with the recent trend of citizens participating in disaster response by creating applications gathering and exchanging knowledge and information to official maps, which are open accessible for the public (Turoff, Hiltz, Bañuls, & Van den Eede, 2013).

With the advanced technology and its use, we've come to large volumes of data spread over multiple datasets. The variety of information management rules and software makes it hard for the humanitarian actors to use the data. Locating health facilities in disaster areas is a good example of this challenge. In terms of health facilities, many databases are available. Some, such as Open Street Map (OSM), are easily accessible but still largely incomplete, while others have more exact data but are not easily shared outside of the organizations which have gathered them. All these datasets complement each other in terms of geographical coverage and in terms of the information they contain, however they are almost never readily available in a consolidated, freely accessible way. This can seriously hamper initial relief efforts in emergencies (Saameli, Kalubi, Herringer, Sutton, & Roodenbeke, 2016).

To address these issues healthsites.io, the International Committee of the Red-Cross (ICRC) and the International Hospital Federation (IHF) have joined their competences and networks in order to create a free, curated, global source of healthcare location data.

1.1 Volunteering and…

1.1.1 …Crisis Mappers

The world is facing an increasing number of complex natural or man-made humanitarian crises. In order to respond to these growing challenges, humanitarian actors are deploying more and more innovative technologies and approaches to support relief aid (Haselkorn, 2009). Volunteer crisis

[3] Faculty of Geodesy University of Zagreb, e-mail: dodobas@geof.hr
[4] Faculty of Geodesy University of Zagreb, e-mail: fgrubisic@geof.hr

mappers are a recent phenomenon driven by technological advances, online communities, and volunteerism. With the rise of interactive websites, social media, and online mapping tools, individuals across the world are able to participate in collecting data in response to a crisis (Resor, 2015).

1.1.2 …Geographic Information

Geo-information are an example where humanitarian practices have dramatically evolved in the recent years with the emergence of the phenomena called by (Goodchild, 2007) Volunteered Geographic Information (VGI), or more generally by (Burns, 2014) "digital humanitarianism". The immediacy of locational information requirements and importance of data currency for natural disaster events highlights the value of volunteered geographic information (VGI) in all stages of disaster management, including prevention, preparation, response, and recovery. VGI technologies enable rapid sharing of diverse geographic information for disaster management at a fraction of the resource costs associated with traditional data collection and dissemination, but they also present new challenges. These include a lack of data quality assurance and issues surrounding data management, liability, security, and the digital divide (Haworth & Bruce, 2015).

2. THE IDEA

The idea behind this open development initiative, called The Global Healthsites Mapping Project, is to create an online map of every health facility in the world and make the details of each location easily accessible. No such global data currently exists. Currently health care location data is either incomplete or not available in an accessible open manner. Healthsites.io aim to provide a single point of reference for healthcare workers, aid agencies, contingency planners, government agencies and citizens who need access to a highly curated global dataset of healthcare facilities. By simply clicking on the map users can discover what healthcare facilities exist at any global location and find out what resources exist at any individual healthcare facility.

In order to achieve this goal, the project has developed a specific master data management methodology which the healthsites.io team are in the process of implementing.

2.1 Vision

When a natural disaster or disease outbreak occurs there is a rush to establish accurate health care location data that can be used to support people on the ground. This has been demonstrated by events such as the Haiti earthquake and the Ebola epidemic in West Africa. As a result, valuable time is wasted establishing accurate and accessible data. Healthsites.io establishes this data and the tools necessary to upload, manage and make the data easily accessible (Healthsites.io, n.d.).

The Healthsites Map will be of great value to members of the public, to health ministries and healthcare providers, and will be invaluable to first responders in emergency relief, disease epidemic and crisis situations.

Healthsites.io is based on four pillars:

1. **Global Scope** – The Global Healthsites Mapping Project is an initiative to create an online map of every health facility in the world and make the details of each location easily accessible,
2. **Open Data Collaboration** – Through collaborations with users, trusted partners and OpenStreetMap, capture and validate the location and contact details of every facility and make this data freely available under an Open Data License (ODBL),
3. **Accessible** – Make the data accessible over the Internet through an API and other formats such as GeoJSON, Shape files, KML, CSV,
4. **Focus on health care location data** – Design philosophy is the long term curation and validation of health care location data. The healtsites.io map will enable users to discover what healthcare facilities exist at any global location and the associated services and resources.

2.2 Why build a new database/service?

A. **Global focus**: All the data for the world in one place. Existing tools tend to have a narrow focus on specific region.
B. **Clean extensible code base**: Should not be hard to pick up and go with the code
C. **Rapid deployment**: Should be able to spin up an instance with minimal hassle
D. **Highly validated data**: Have great focus on providing quality measures on Healthsites.io data
E. **Community validated**: Enabling community to validate data based on community/social tools
F. **Clean simple design**: Build a tool that is highly accessible for users and developers with a Zen like minimalism, narrowly focused on implementing Healthsites.io vision.
G. **Standards compliant**

3. HEALTHSITES.IO APPROACH

3.1 Datasets integration

OSM, with its large number of entities, worldwide coverage and large community, was chosen as the main dataset of the project. Also, some databases from trusted partners were chosen. All this data together gathered and represented more than 170 000 health localities worldwide. In order to integrate all these datasets, which have different structures and values, and a data model, *Entity–Attribute–Value* model, based and inspired by the OSM data model, was chosen. It enables users to store information about anything. A FullTextSearch index has been implemented to enable searching for textual data. The importing process is based on a similar approach; a user needs to

define an attribute mapping file for Attributes that are going to be imported. These attributes are defined as either Mandatory (Table 1), Core (Table 2) or Other (Table 3).

Table 1. Mandatory Attributes

No	Validatio	Healthsi	value
1.	mandator	uuid	unique id
2.	mandator	date-	date time
3.	mandator	name	Name of

No	Validatio	Healthsi	value
4.	mandator	geom	lat/long
5.	mandator	source	Name and

Table 2. Core Attributes

No.	validation	Healthsite	value
5.	core	physical-	address
6.	core	contact-number	number including international dialing code (+250 252
7.	core	nature-of-	clinic without beds, clinic with beds, first referral
8.	core	scope-of-	all type of services, specialized care, general acute care,
9.	core	operation (2)	24/24 &7/7 ; open only during business hours ; other
10.	core	inpatient-	Number of full time beds, number of part time beds
11.	core	ancillary-	Operating theater, laboratory, imaging equipment,
12.	core	activities	medicine and medical specialties, surgery and surgical
13.	core	staff (4)	X full time equivalent doctors and Y full time equivalent
14.	core	ownership	public, private not for profit, private commercial
15.	core	raw-data-	a link to the raw data file (.csv ect)

Table 3. Other Attributes

validatio	Healthsite
other	contact_perso
other	phone
other	mobile
other	email

validatio	Healthsite
other	url
other	country
other	Status
other	Verified

validatio	Healthsite
other	Date entered
other	Centre ID
other	Org
other	Capacity (N)

validatio	Healthsite
other	Equipment
other	Staff (N)
other	Doctors
other	Nurses
other	Isolation
other	PPE
other	Burial Site

validatio	Healthsite
other	Water
other	Power
other	Comm.
other	Date open
other	Date closed
other	Province
other	Adm1 ID

validatio	Healthsite
other	District
other	Adm2 ID
other	Place
other	Adm3
other	Location in
other	Notes

However, this flexibility of storing information represented a risk potentially hampering the goal of having a curated database with easily accessible information. In order to address this issue, 15 core attributes (Table 1 + Table 2) that are relevant for both the public and health professionals were defined by the International Hospital Federation. For most of these core attributes, defined values were set to enable comparisons. Indeed, an essential question such as "What type of health facilities is it" can vary greatly according to the national or organizational classification system (Saameli, Kalubi, Herringer, Sutton, & Roodenbeke, 2016).

As lot of data is expected, server based clustering approach has been opted. It offers full control of what actually gets clustered, as some search queries might return a lot of data, or some data is only visible for certain users. Clustered rendering minifies amount of data transfer between client and the server, which is nice for areas with limited connectivity or slow 'mobile data plans'.

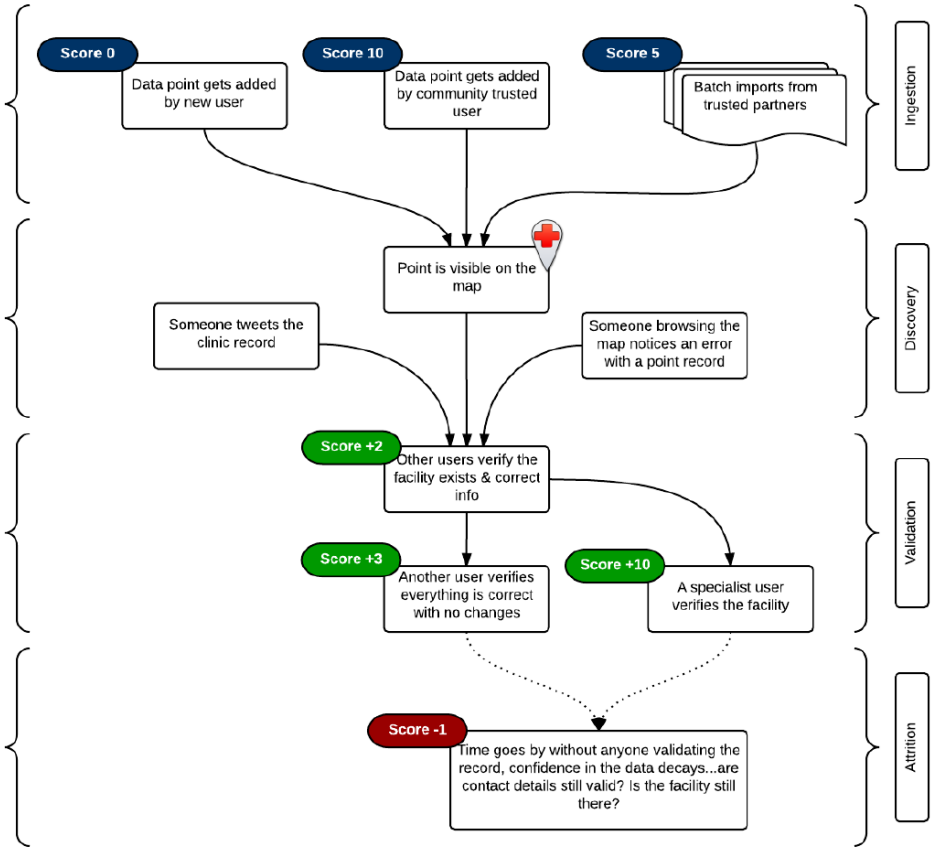
3.2 Validation Process

To be able to achieve reliability of health facilities data, some validation processes have been planned to be added to be project, which vary from simple to more complex ones. Every attribute can have one or more associated validators, with is ranging functionality: "email address should look like an email address' to more complex (which even rely on external services), like: "check if the Locality address is similar to results returned by external geocoding services".

There will also be a validation process based on user reputation (Figure 1) to assess the reliability of data through a Locality Validity Index (LVI). This reputation based process consists of 4 steps. The first step takes place during data integration. Depending on source of the data, the LVI obtains a score ranging from 0 to 10; 0 being for data added by a new user, 10 for data from a community trusted user and 5 if it is a batch of data coming from a Ministry of Health or a healthcare organization. The reliability of the user is based on the monitoring of his activity on the platform and the crosscheck of his activity by other users. Secondly, once the data is displayed on the map of

Healthsite.io, any user can complete missing attributes data and modify or validate existing attributes data through a tweet channel. Thirdly, the more users confirm the validity of information, the higher the LVI becomes. Within this step, if an authoritative user such as a staff of health organization, verifies information, the LVI becomes even higher. Finally, if time goes without anyone validating the record, the LVI progressively decreases (Saameli, Kalubi, Herringer, Sutton, & Roodenbeke, 2016).

Figure 1. Reputation based validation process



3.3 Data Life Cycle

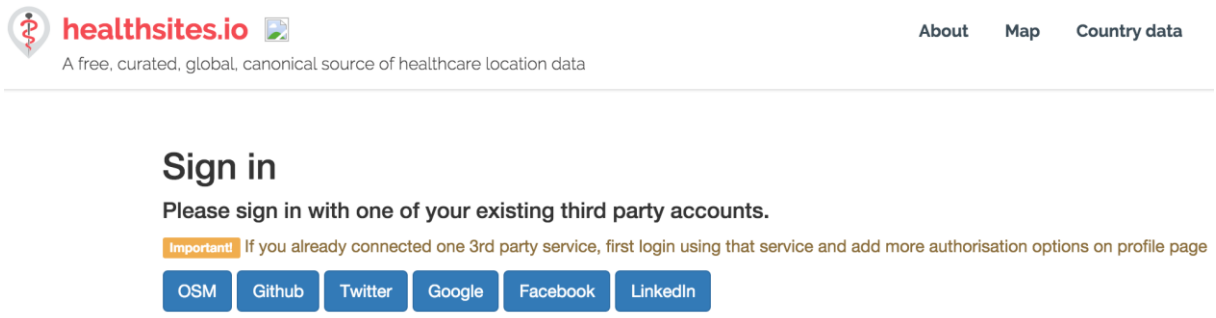
Healthsites.io differentiate themselves from other great open dana projects like OpenStreetMap (OSM), Ushahidi, Sahana etc. by their design philosophy. The focus is on the long term continued curation and validation of dana for a single domain, and then exchange and share dana freely with other initiatiaves in a way that healthistes will import dana from OSM, validate and improve the data, and then push it back to OSM.

There will be a reputation system, where each site is awarded points that speak to the credibility of the given dana at that site. Points will be awarded based on interactions with the site, or lost when no interactions occur for a period od time.

The Healthsite.io project has been designed to have a bottom-up approach of the monitoring and data collection. Indeed, it is OSM community who provides the localisation of health facilities and health workers in the field who provide related attribute information, such as the type of services available. These two sources of data are consolidated in Healthsite.io and data quality verified through the LVI. Once data quality is good enough, data are then sent back to OSM in order to be shared widely through the community of OSM users (Saameli, Kalubi, Herringer, Sutton, & Roodenbeke, 2016).

3.4 Design and User Interface

Figure 2. Collaborative user feedback



Social authentication backend enables users to login using external services like OpenStreetMap OAuth or Google/Facebook/Twitter/LinkedIN. Authenticated users can make changes to existing, or create new, healthcare location data.

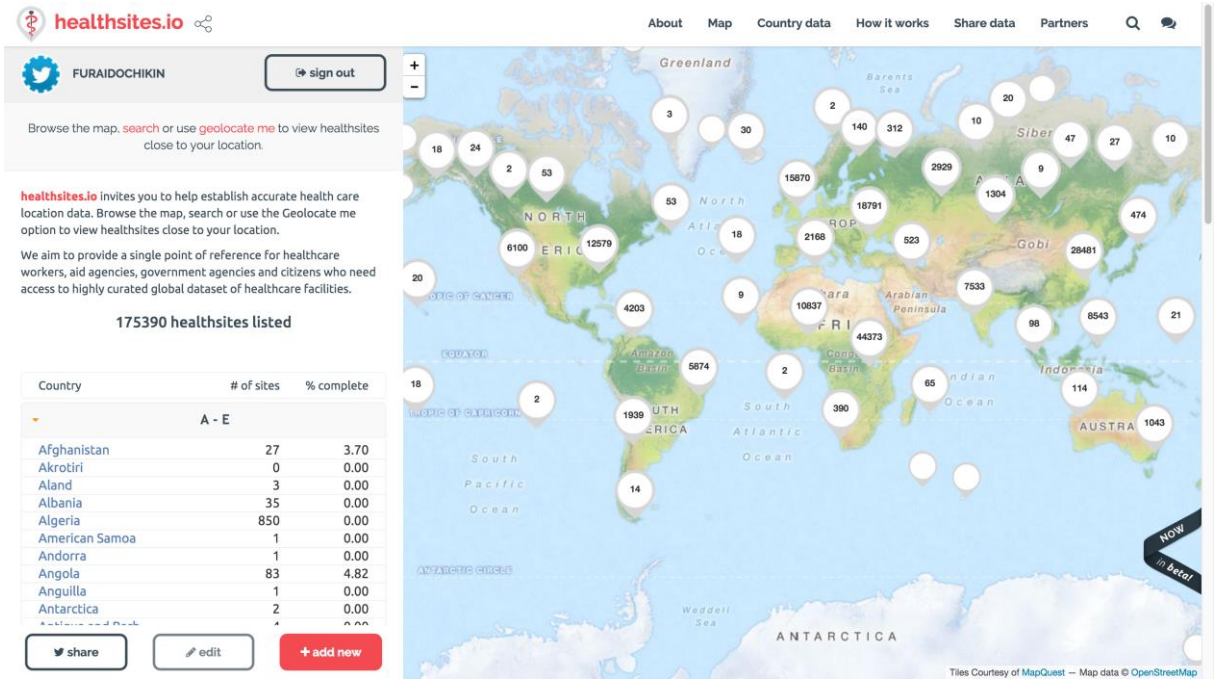


Figure 3. Logged in Interface

Logging in brings you to this interface, divided into two main pieces – the left sidebar and the map. On the left sidebar you can use tools such as search or geolocation request, to browse the map the way you want. Also, the key idea of Healthsites.io is written down and explained. Furthermore, you can see the current number of listed Healthsites, as well as the complete list of countries, along with number of sites in the particular country and its completeness.

When signed in, by clicking on *+add new*, you can easily change the Healthsites attributes, as shown on the picture below (Figure 5). By clicking *create*, you save the changes you have made and update the healthsite. You can also share the newest update on Twitter, with an embedded link (*Share*) and post (Figure 4)

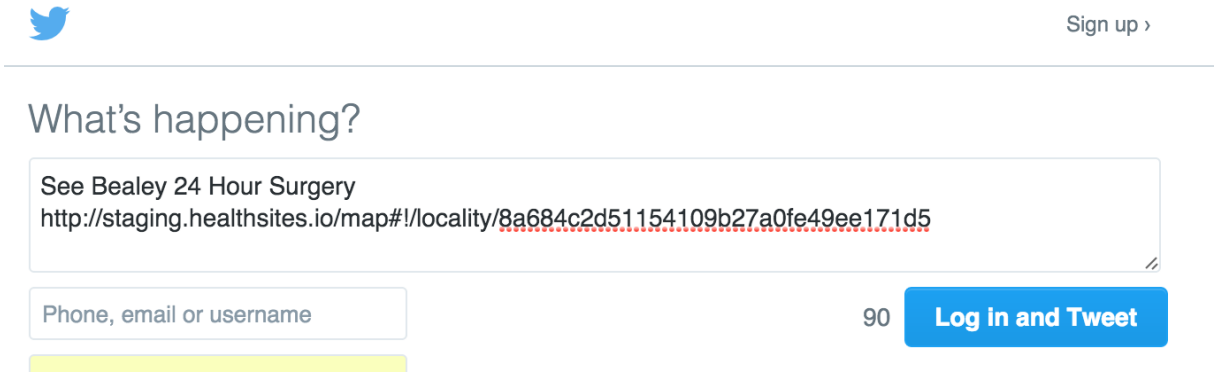


Figure 4. Update Sharing

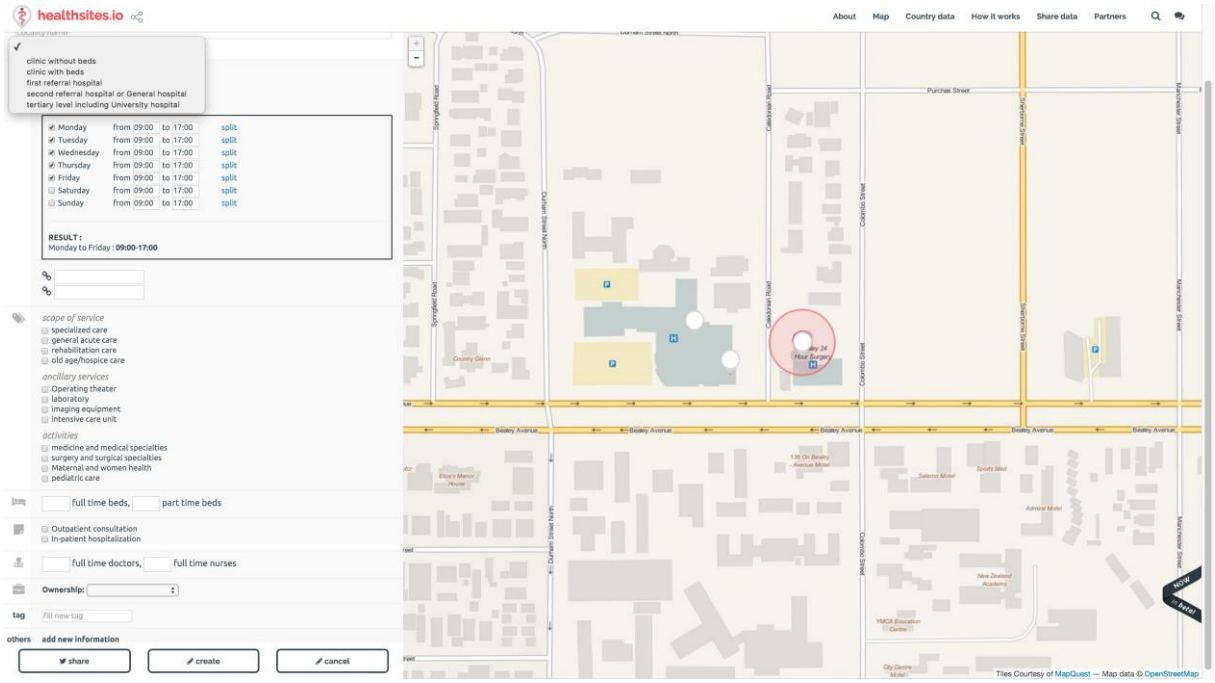


Figure 5. Updating Attributes

Click on a particular hospital marker icon loads you this interface, showing all the relevant details and attributes. When some information is needed, the Heading lights up blinking red, luring your attention (Figure 6).

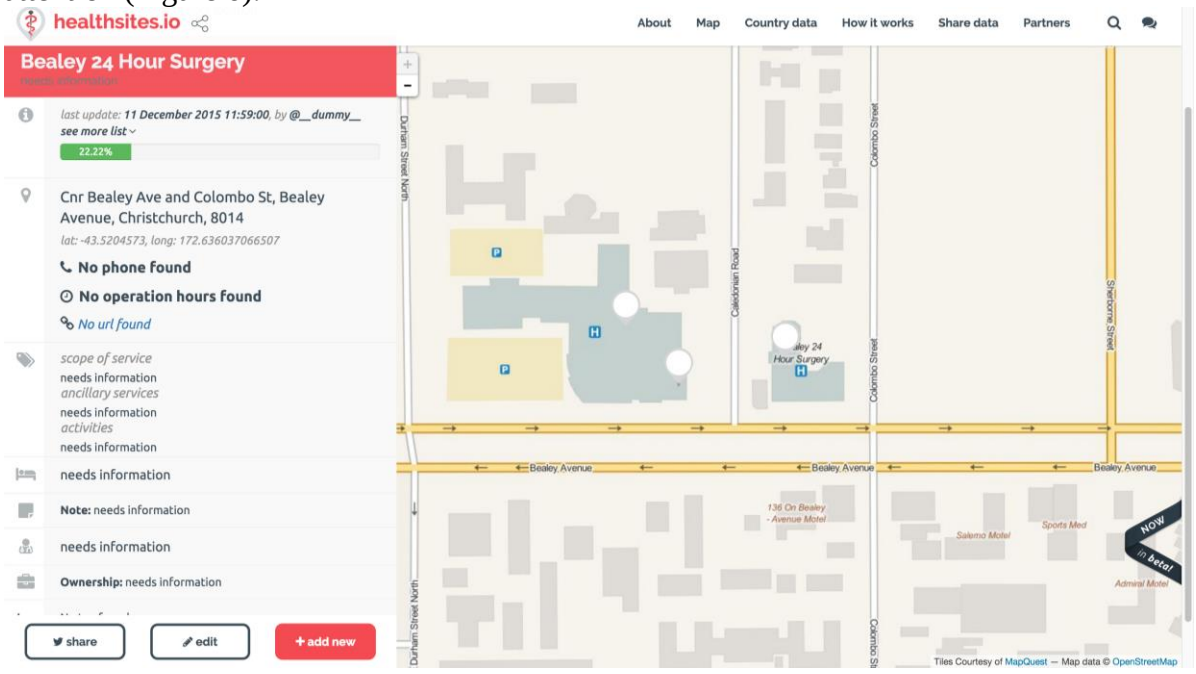


Figure 6. Healthsite Information

The interface is very user friendly, showing you in the progress bar the updating and completeness progress and when the last update was and by whom. Other key information shown in the left sidebar are *Location (Physical Address and lat/long), Contact Number, Working hours, Website, Service, Bed information, Staff, Ownership.*

Figure 7. Country Data Information

When clicking on "Country Data", this interface is loaded (Figure 7). By searching a country name, in this example New Zealand, an interactive bar chart and pie chart are shown, with a total count of Healthsites in that particular country. The data is shown easily and visually pleasant, showing a number of Healthsites per type, varying from Hospitals, Medical Clinics and Orthopedic Clinics in the bar chart; as well as the completeness of attributes in the pie chart, shown in percentage of Complete, Partial and Basic.

3.5 Licenses

3.5.1 **Data -** Open Database License

The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Database while maintaining this same freedom for others. Many databases are covered by copyright, and therefore this document licenses these rights. Some jurisdictions, mainly in the European Union, have specific rights that cover databases, and so the ODbL addresses these rights, too. Finally, the ODbL is also an agreement in contract for users of this Database to act in certain ways in return for accessing this Database.

3.5.2 **Code -** Free BSD License

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

3.5.3 Content - Attribution 4.0 International

You are free to:

- **Share** — copy and redistribute the material in any medium or format
- **Adapt** — remix, transform, and build upon the material

for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

- **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- **No additional restrictions** — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

No additional restrictions — You may not apply legal terms or **technological measures** that legally restrict others from doing anything the license permits.

4. CONCLUSION

This project aims to democratize access to health geodata. By enhancing the quality and accessibility of geodata, this project offers several opportunities to achieve social, health and humanitarian impacts in areas where data is scarce. A better knowledge of locations of health facilities and associated services can help to map and identify population underserved by healthcare services (Muenoz & Källestål, 2012) (Blanford, Kumar, Luo, & MacEachren, 2012). Such information can contribute to develop and advocate better evidence-based health policy. It can also help healthcare organizations to better plan their activities, vaccination campaigns based on the correlation between vaccination rate and distance to primary healthcare center (Al-Taiar, Clark, Longenecker, & Whitty, 2010). For relief workers, knowing where the health facilities and associated health service are located are crucial information to know to prepare contingency plans, but also to response promptly and meaningfully during the emergency.

Depending on the future contributions of VGI communities and its development, the goal and the capacity of the project to meet its aim – to build global, complete and high quality database of world health relies on it. However, when comparing with other similar VGI initiative, this project has the advantage of having the endorsement of major actors such as Humanitarian OSM team (HOT), the International Committee of the Red Cross (ICRC) and International Hospital Federation (IHF). This support enables a strong promotion among digital humanitarians and health experts.

Finally, the model of Healthsites.io of taking OSM data to enhance its attribute by health experts before exporting back completed and validated data to OSM is a truly innovative approach. It fosters bidirectional exchange between health experts and VGI communities. In addition, it also a new approach to quality issue of attribute data in OSM. If this model proves to be successful, it could be replicated in other essential domain, such water services, where accessibility to detailed and reliable data is much needed for humanitarian operations. This approach would be a new way to maintain updated comprehensive thematic database without having to spend resources usually required for such tasks (Saameli, Kalubi, Herringer, Sutton, & Roodenbeke, 2016).

REFERENCES

Turoff, M., Hiltz, S., Bañuls, V., & Van den Eede, G. (2013). Multiple Perspectives on Planning for Emergencies: An Introduction to the Special Issue on Planning and Foresight for Emergency Management. *Technological Forecasting and Social Change*.

Saameli, R., Kalubi, D., Herringer, M., Sutton, T., & Roodenbeke, E. d. (2016). Healthsites.io: The Global Healthsites Mapping Project. 6.

Dijkstra, P., Grubišić, F., Leitz, S., Kagerah, S., & Unger, E.-M. (2015). FIG YSN Mapping Response - Contributions for Nepal. *FIG-ISPRS International Workshop Post 2015 Nepal Earthquake*, (p. 8). Kathmandu.

Healthsites.io. (n.d.). *Github.com/healthsites*. Retrieved 2 9, 2016, from Vision: https://github.com/healthsites/healthsites/wiki/Vision

Resor, E. (2015, 12 10). The Neo-Humanitarians: Assessing the Credibility of Organized Volunteer Crisis Mappers. *Policy and Internet*.

Haworth, B., & Bruce, E. (2015, 5). A Review of Volunteered Geographic Information for Disaster Management. *Geography Compass, 9*(5), 237-250.

Goodchild, M. (2007). Citizens as sensors: The world of volunteered geography. *GeoJorunal, 69*(4), 211-221.

Burns, R. (2014). Rethinking big data in digital humanitarianism: practices, epistemologies, and social relations. *GeoJournal, 80*(4), 477-490.

Haselkorn, M. &. (2009). The Role of Information and Communication in the Context of Humanitarian Service. *IEEE Transactions on Professional Communication, 52*(4), 325-328.

Muenoz, U., & Källestål, C. (2012). Geographical accessibility and spatial coverage modeling of the primary health care network in the Western Province of Rwanda. *International Journal of Health Geographics, 40*(11).

Blanford, J., Kumar, S., Luo, W., & MacEachren, M. (2012). It's a long, long walk: accessibility to hospitals, maternity and integrated health centers in Niger. *International Journal of Health Geographics, 24*(11).

Al-Taiar, A., Clark, A., Longenecker, J., & Whitty, C. (2010). Physical accessibility and utilization of health services in Yemen. *International Journal of Health Geographics, 38*(9).

BIOGRAPHICAL NOTES

Franka Grubišić is an undergraduate student on Faculty of Geodesy University of Zagreb. She has been Head of IT Section of the Student Council since 2014, and from October 2015 she is a Student Representative at the Council of Technical Sciences of University of Zagreb and a Teaching Fellow on the Faculty of Civil. Her project "*New Sightings of Geodesy*" has won the CLGE 2014 Student Contest and she has been awarded with a Scholarship from the Institute for Energy Sources Croatia.

Dražen Odobašić is a Researcher Fellow on Faculty of Geodesy University of Zagreb, currently on an ongoing PhD, where he teaches *Databases, Geospatial Databases, Scripting Languages in Geodesy and Geoinformatics, Open Geoinformation* etc. His interests include open source, FOSS, geospatial analysis and so on. Aside working as a Research Fellow, Odobašić has a number of other projects, one of them being Healthsites.io, in which he is a *Geospatial Analyst using FLOSS.*