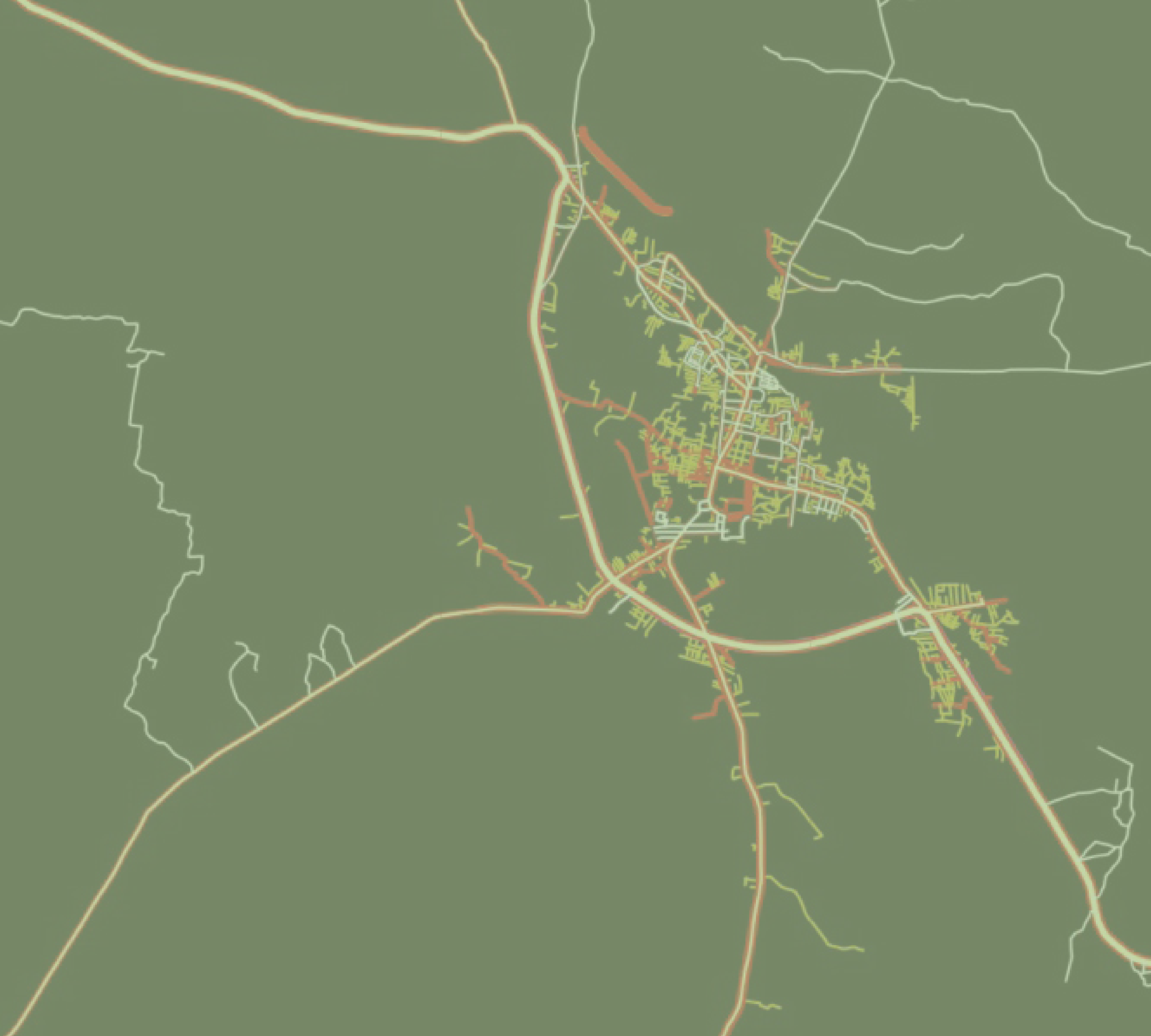
Knowledge Product

Online Tool for City-Level Slum Free Planning

British High Commission New Delhi

SCI-FI
SCALING CITY INSTITUTIONS FOR INDIA
LAND HOUSING PLANNING

CENTRE FOR POLICY RESEARCH



Contents

List of Figure

1. Background

The use of Geographic Information Systems (GIS) was an essential component in the planning and execution of the JAGA Mission, also known as the Odisha Liveable Habitats Mission. In the course of the Mission's execution, a substantial amount of spatial data was generated and used for accurate mapping of slums, identification of beneficiaries, calculating the exact extent of land occupied by slum dwellers, etc. The Government of Odisha wished to further update, add to, standardize and analyze the JAGA Mission data sets for spatial analysis of various kinds. It also wishes to empower urban local governments and community organizations (slum dwellers associations) by allowing them to use and get command over their data. An open-source spatial analysis platform was developed to support the Government of Odisha (GoO) in this task and piloted in Dhenkanal.

The concept and the genesis of this tool stem from the Toolkit originally built for Open and Sustainable City Planning and Analysis, developed in cooperation between the Digital City Science of the HafenCity University Hamburg (HCU) and Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH (GIZ) in India and Ecuador. The tool's objective is to support integrated and participatory urban planning processes, enabling dialogue between governments and citizens and the exchange of knowledge and data between government departments. The main functions of the tool are to visualize and analyze complex urban data, jointly among local practitioners and with citizens[1].

[1] https://github.com/citysciencelab/open-city-toolkit/wiki/1.-Toolkit-for-Open-and-Sustainable-City-Planning-and-Analysis-%E2%80%90-User-manual

Government plays a crucial role in preparing cities to be climate resilient by quickly understanding and responding to climate change impacts. Achievement of improved response time will require data insights to ensure informed decision-making—at different scales of governance (local, state, national or global). Currently, there is a proliferation of new technologies to help us monitor our habitats - satellite imagery, sensor data about temperature, precipitation, air quality, wind, soil conditions, and more. The issue is not about the lack of data but the capacity of the government to harness this data to drive insights for decision-making for climate change. These insights help us better manage resources, understand risks, predict changes, and respond to disasters as we meet the challenge of climate change.

Geospatial tools like the City App provide government departments and agencies with a centralised system to collect, process, and analyse these data into decision-making contexts. They offer insights that are helpful for a wide range of departments and use cases and can be applied in spatial and temporal scales that are important for governments to take action. It can enable government agencies to fulfill their mandates by drawing invaluable insights into changing urban landscapes, about physical and natural assets at risk, designing policies that take into account these ground realities.

As in the case of Odisha, only slum households on untenable lands were identified and relocated, therefore minimizing the relocation of slums. Use cases may also include climate change impact themselves. These include monitoring, predicting, and analysing the risks of extreme weather events like floods, extreme heat, storms, drought, and other climate hazards and may also include tracking changes in the environment, like disease vectors, water availability and quality, soil conditions, air pollution, and more.

2. Tool built around the reality of the JAGA Mission

2.1 Creation of data

Spatial data generation as part of the JAGA Mission started in 2017, where a vast amount of data was created in about seven to eight months. It covered 1725 slums all over the state, with data of 1 70,000 households. The data was enormous in its sheer breadth and contained 2-2.5 cm resolution drone survey data. These were heavy files, digitized along with the household data. This data was being collected by multiple agencies, through the Tata Trusts, partner to the GoO, and Omidyar Network. The mapping exercise focused on the Land Rights Act parameters, based on which the correct beneficiaries would be identified. Given the scale of operation across 109 cities across Odisha, this method was considered the most efficient, accurate, and transparent.

2.2 Data creation is not sufficient; the need for ownership and management of data.

There were multiple agencies involved in the data collection. Some have been involved in the drone survey, some had been involved in the digitization, and some had been involved in the initial household data linking with the GIS data. Data were collected and stored in different vendors' servers; it was not managed or owned by any. The data was created for the government for implementation of the Mission. The government had to own the data, manage and use it. So there was this vast repository of digital spatial data created as part of the JAGA Mission by large GIS companies for the GoO. It was the question of the government's capacity and its partner's capacity to handle geospatial data. The software (mostly expensive, proprietary software) and the skills belonged to the private companies contracted to prepare the digital data sets. Once the contract periods ended, the digital data was handed over to the government, but not the software or the operational skills.

The process was clumsy (spend money to create digital datasets and then work with paper maps!), but it also meant that for every spatial decision, the revenue team had to deploy teams to undertake field surveys cross-verify the paper maps. Not a desirable process in the days of the pandemic.

However, the spatial data increased the efficiency of resource allocation by the revenue department by targeting field verification only to specific problem cases identified by the digital tool. This meant that teams had to be deployed only when necessary, thus substantially reducing unnecessary risks to government staff during the pandemic. This further helped convince the government about the potential of the tool.

2.3 Technology as enabler

During the next couple of years, the spatial data was not being used, and the GoO resorted to the old manual mode of operations. So, when CPR signed an MoU with the H&UDD, GoO in September 2020, several aspects were identified where technology could be an enabler and make things work more efficiently, as it did in Phase 1 of the JAGA Mission when technology was used effectively and resulted in a massive amount of data creation within months (drone survey and GIS-enabled household survey). The use of technology could again enable JAGA Mission implementation to be more efficient and seamless.

Dhenkanal Municipality initially wished to use the tool primarily to familiarise its staff with digital spatial data. Their interest lay mainly in understanding specific problems they face concerning issues of the untenability of land parcels. Contrary to what we had initially expected, there is no hesitation or resistance to using the tool. Even the revenue experts, who generally tend to be conservative about technique, realised the potential of digital data. However, they provided valuable feedback regarding integrating slum level layers to revenue land parcel layers. The revenue layer styles were altered based on the input, and the boundaries were turned into bands. This helped clarify many of the concerns that the Municipality and the revenue officials had concerning identifying slum households that lie on untenable land parcels.

2.4 The rationale for technology choice-Open-source software

Data management cannot be outsourced. It has to be done internally within the government system. So, a tool was needed to manage and effectively use the data. However, the government needed to be convinced about the possibilities and the tool's power to aid them in their day–to–day operations. Moreover, the government might not invest in high-end proprietary software like ESRI Arc GIS, etc. The tool also had to be flexible and dynamic enough to cater to the ever-changing needs of the government. The choice of the tools focused on the best way of managing this kind of data (JAGA Mission data focus) and how the government could take ownership of its data and make sense of it. A whole range of open-source tools would be imperative.

3. About the Tool

- An open source digital tool for spatial analytics.
- Collect, clean, standardise, analyse and visualise spatial data for decision making.
- Consists of a distinct front-end (for visualisation and consultation) and back-end (for data storage, processing and GIS engine)

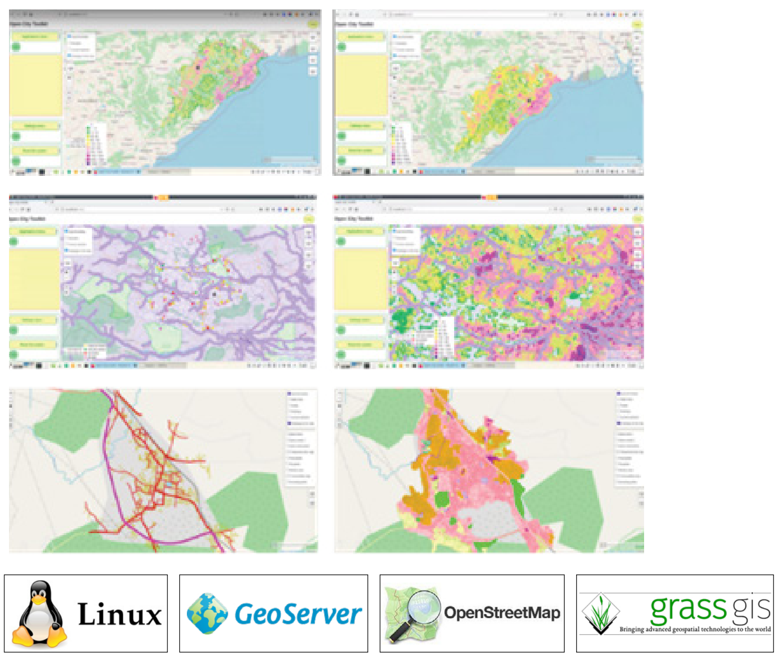


Figure 1: The geo spatial analytic tool, developed by the SCI-FI CPR team

The use of Geographic Information Systems (GIS) was an essential component in the planning and execution of the JAGA Mission. However, while ample spatial data was generated and used for accurate mapping of slums, identification of beneficiaries, calculating the exact extent of land occupied by the slum households, etc., there was a lack of adequate use of available data and spatial analytics in making policies more effective and relevant. Accordingly, CPR designed an open-source, web-based spatial analysis platform to support the GoO in this task and piloted it in the city of Dhenkanal by enabling the ULBs to identify slums, analyse

The interaction with the Slum Dwellers Association of Kathagada Parbatia Sahi and a women's self-help group belonging to the same slum turned out to be the liveliest one. Surprisingly, the slum dwellers appeared to quickly comprehend digital maps. The community's interest lay primarily in understanding and documenting what is going on in their settlements and their environs. It was evident that the community members were more comfortable with smartphones than laptops. Also, the access to smartphones was much more than access to computers and laptops, despite some young members of the slum attending computer education courses. Based on this feedback, an android version of the tool was developed.

the infrastructure gap and provide necessary interventions. Several consultations were held, and updates were made in the tool based on the feedback. An android version of the tool was also developed.

The tool intended to combine different types of digital data and process them on a platform to facilitate both the testing and visualisation of existing models' outcomes and help generate newer and more fine-tuned models. The key objective of the tool was that it had to be a flexible system, a presentation bridge where the data could be presented. It enabled the GoO to take care of its data and use it efficiently for decision-making. This would foster more transparent and evidence-based planning, multi-stakeholder decision-making, and driving digitalisation in different sectors of local city governments. Furthermore, it would help public administrations overcome the sectoral divide by efficiently and comprehensively sharing spatial data. The tool offers the following:

- A simple visualization platform that brings together spatial information and complex analytics that is easily comprehendible by decision-makers and other non-technical stakeholders
- Allows for different kinds of analyses through data integration - cross-layering or overlaying information
- Stakeholders without GIS expertise can also conduct their analyses due to the simplified nature of the tool.
- Facilitates direct interactive discussion between experts and non-experts, leading to evidence-based multi-stakeholder decision making
- Participants can help bridge a data availability gap by providing their valuable local knowledge of the city and area in question. This could result in complementing already available data.
- Empowers both the government and the citizens to work as partners and interact together in decisions that affect their lives
- Applications are not limited to urban planning and can be extended and used by different domains like energy, health education, etc.

The key aspects of the tool are highlighted in **Annexure 1**.

4. The process of tool development

The tool development process is key to the usage and relevance of the tool.

- **Data convergence:** All necessary spatial and statistical data (high-resolution drone imagery of slums; Jaga GIS database for the pilot city; Jaga Mission 'Urban Slum Household Area (USHA)' survey data for the pilot city; as well as other related survey data covering infrastructure characteristics of the pilot city and region) were collected from the relevant Departments/Mission offices.
- **Screening and creation of metadata file**: Preliminary screening of all statistical and GIS data for the pilot city; metadata were recreated.
- **Identification of data inconsistencies and gaps** – Internal inconsistencies and errors in the data and conflicts between the spatial and statistical databases were identified and resolved.
- **Data cleaning, structuring, and organization**: All the layers used in this tool were cleaned and organized so that the data files can be identified by new users intuitively, along with relevant metadata. If one has an administrator login, the data can be updated anytime and viewed using any open-source GIS software like Q GIS.
- **A collaborative model of software development:** Unlike the usual outsourcing model of software development, this was a product of a process that was largely internally driven by the knowledge partner CPR. The software modules were built based on continuous interaction with the in-house technical partners, the government, and GIS experts.

Challenges of working with large databases created through government schemes in India

- **Missing or incomplete metadata files**
- **Major conflicts between the spatial and statistical databases**

The Department of Housing & Urban Development (HUDD), being the state level nodal agency for implementing Jaga Mission, wanted to use the tool mainly for state-level consolidation of spatial data; discovered errors in the data which was cleaned and standardised, and then used for resolving pending cases of in-situ land rights; identifying suitable parcels enabling slum proofing and state-level training and orientation activities. Hence, a state-level base template was added on which the detailed city and slum level data can be added incrementally as data becomes available.

- **Incremental development of the tool use case functions:** A critical spatial issue that the government is dealing with is identified, and spatial data analysis may be pertinent for a quick and efficient resolution; a solution script is created iteratively. It is created like a function first to resolve the problem at hand. Then, it is converted to a module only if the problem is universal for the state and will be replicated. In many cases, even scripting was not required; just visualization of relevant thematic layers was adequate.

5 The enablers for the tool

This section will describe how the process of tool development stemmed from the existing initiatives and continuous engagement and understanding of the stakeholders' needs. This led to the development of the tool designed by and for the government, contributing to the success of the tool use and its scalability.

5.1 Partnerships with the Government

CPR and Janaagraha were the two key government partners in the implementation of the JAGA Mission. CPR's contribution was particularly related to the technical aspects of policymaking, preparation of SOPs for the slum upgrading phase, and advocating the use of technology in decision-making. Whenever the government of Odisha gets into a partnership with organisations, they are expected to set up field office/s within the state. They have an operating unit in the JAGA Mission Office in Bhubaneshwar. The key is working near the action on the ground, both metaphorically and geographically. So, positioning internally within the government system was a strong enabler and facilitator.

5.2 Quick resolution through demonstration

The demonstration of the potential of the tool led to the following outcomes.

- **It also brought in an immediate behavior change, with senior officers in H&UD requesting all available GIS maps to be pulled up anytime a discussion involving spatial decision-making starts.**

1. Confidence to take ownership of data
2. Understanding the limitations of spreadsheets and the power of spatial tools
3. Optimising resource allocation and fostering inter-department collaboration
4. Aiding the process of slum relocation

5.2.1 Confidence to take ownership of data

When CPR started working with the GoO in 2020, it was aware of the potential of the Open-Source Toolkit, developed by GIZ and HCU, and had experience with using this tool with Odisha data. Over time, CPR demonstrated how the tool works and helped resolve issues very efficiently. The tool was useful, but there are also issues in the realm of geometric or geographical analysis and beyond the capacity of excel, and spatial visualization and analysis become imperative, for

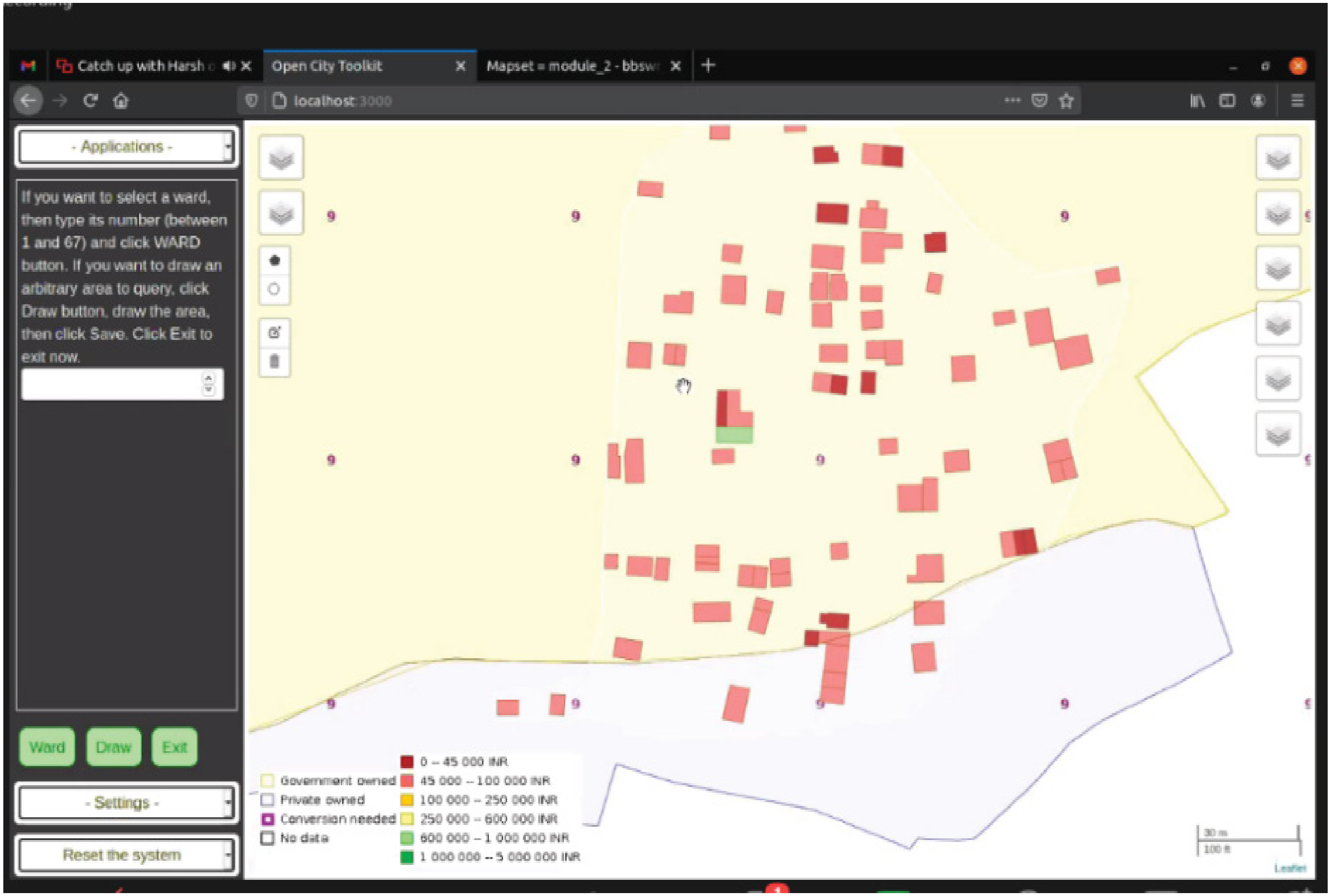


Figure 2: Simple overlay of slum households, slum land, and revenue land resolved the duplication of Land Rights Certificates.

example, calculating intersecting areas within boundaries of slum and revenue lands. These calculations can be done instantly through simple functions of the tool and thus help decide policies and help negotiate transparently with the stakeholders.

The digital tool gave the confidence to GoO to take over the data from the GIS companies. Until now, the government was doing the tasks, but the data was in custody and servers of the large GIS companies, who had helped create the data. CPR offered to be the "ally" of the GoO, their team driving the demonstration of the potential of the tool, using the data that was already created, and also other data that is available statewide, like open street maps and data from Odisha Space Applications Centre (ORSAC). This gave the government enough confidence to ask the GIS companies to hand over the data to them and gain ownership. The CPR team initially used the tool for simple communication of issues on the ground. The Principal Secretary for the State provided a quick overview of the spatial dimensions of an issue and a pan-state spatial overview. It offered the spatial perspective over and above the numbers and attributes. It saved to travel and back and forth communication, and from Bhubaneshwar, the PS could see the spatial context without being there physically. Just the power of visualization of the evidence brought things into perspective and enabled evidence-based decision-making and efficiency in decision-making. This has also enabled the usage of the JAGA slum data in the city context when other layers like open street maps, where you can see water bodies, road networks, etc., were added. The administrators could use the tool through any open-source geospatial software like QGIS to quickly communicate issues on the ground efficiently. The tool could demonstrate the problem and offer a solution based on the ground reality. And this tool, of course, if it runs on a web server, can operate on a mega scale.

Additionally, adding different data layers at multiple levels enhances the synergy between the state, district, and city. Even slum dwellers associations can visualize and have their stories heard at multiple levels. Supplemental use cases can be derived, as JAGA Mission data is layered with other available data-sets, from different departments and levels.

Although, in its pilot phase, the digital tool contained detailed data of only the city of Dhenkanal, the statewide geographic layer were incorporated later, ensuring that any new data from anywhere in the state of Odisha could be easily incorporated into the system depending on need. This was done by combining geographic data of slums from JAGA Mission with open data such as OpenStreetMap (OSM). This technique is of immense use in a geographically spread-out government scheme such as JAGA.

5.2.2 Understanding the limitations of spreadsheets and the power of spatial tools

- **To Relocate or Not to Relocate**

While the primary focus of JAGA (The Odisha Liveable Habitat) Mission is on providing in-situ land rights to slum dwellers, there are some extreme cases where relocation may be necessary. These are slums located on highly objectionable and hazardous land parcels, such as flood-prone zones, low-lying areas, sites close to waste dumping sites, etc. In January 2021, a request came from the District Collector of Bolangir district to assist in the relocation of nine slums which were located along the Laxmi Jor canal that passes through the city of Bolangir. The banks of this canal get flooded almost every year during the monsoon, inundating the slum settlements located on their banks and causing great distress to the residents. The relocation process involved 1724 families living in nine slums. However, a quick analysis of the situation using the digital spatial tool revealed that only five out of the nine slums were located along the canal. Four slums were not even located close to a water body, let alone be vulnerable to flooding by the Laxmi Jor canal. In all probability, the error happened due to the prevalent tendency in the government to keep working with spreadsheets and not cross-verify the locations using the digital spatial data prepared by JAGA Mission.

LOCATION OF TARGETED SLUMS AND HEW HABITAT SITE IN BOLANGIR

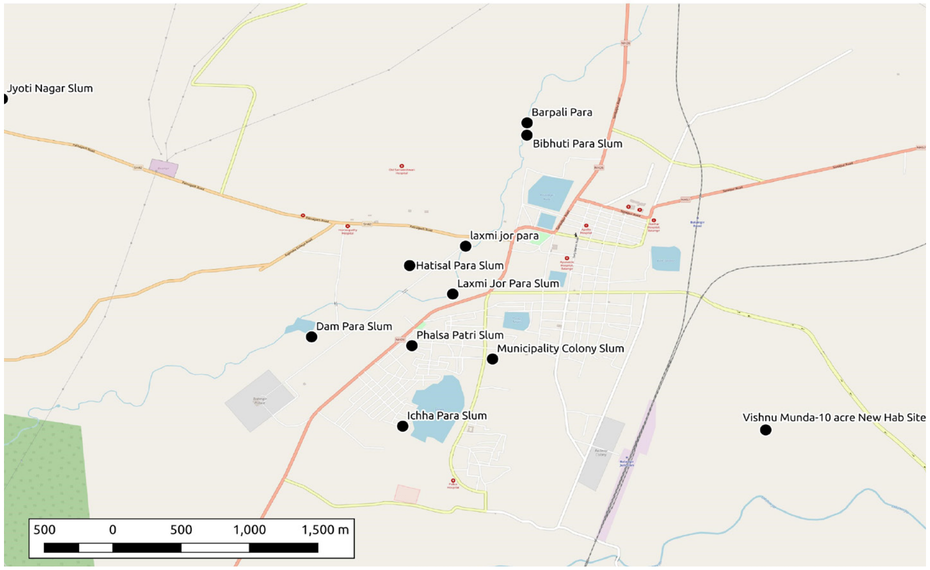


Figure 3: Screenshot of the digital tool showing the slums proposed to be relocated in Bolangir city

Figure 5: Nanda, the boy with the pink bicycle of Hatisal Pada slum, Bolangir – his family need not worry about relocation anymore.

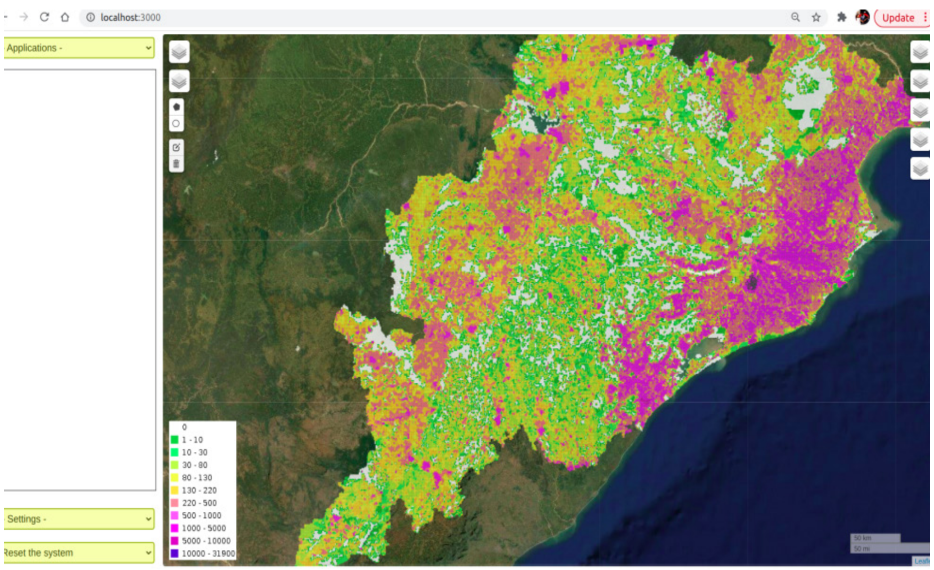


Figure 4: Screenshot of the digital tool showing state-level base layer

Figure 6: Slums along Laxmi Jor Canal, Bolangir city

It took the Principal Secretary of the Housing & Urban Development Department a few seconds to realise this fact – it was obvious just by taking a look at the GIS map displayed by the digital spatial tool. A simple process of visualisation enabled by the digital tool prevented four slums with 468 families from being relocated for no reason at all.

5.2.3 Optimising resource allocation and fostering inter-department collaboration

Interaction with Revenue & Disaster Management, the custodians of all government land in Odisha, is a vital component of JAGA Mission. Once the spatial data of slums of the state was created through UAV-based mapping, it was important to see what kind of land parcels the slums were located on. For this, it was necessary to overlay the JAGA slum maps with the digital maps of the revenue department, which show the land use and ownership details of each parcel of land in the state. The land records maintained by the Revenue department is a world of its containing lists of thousands of land-use categories (for which the Persian word "Qism" is still used just like in the British and Mughal times), which are catalogued under various "khatiyans" (or "khatas" – books) containing unique plot numbers. As slums are informal settlements, they may have come upon land that comprises multiple land parcels of different categories and uses, which may require re-classification of use and interaction with other Departments and committees (e.g., Forest & Environment Department, Endowment commissions for religious land; Department of Railways; private land parcels owned by erstwhile kings, etc.). However, despite all the digital data of these individual components being available, the government was still compelled to work with paper print-outs and spreadsheets. They neither had the necessary software nor the skill set to work with digital spatial data.

The digital spatial tool enabled the overlaying and visualisation of the government's spatial data and removed bottlenecks created by working with paper maps only. Moreover, since the tool is built as open-source software with a general public license (the MIT license), there were no unnecessary expenses and restrictions that necessarily accompanied the use of proprietary software. A classic example of that was when it was argued using the digital spatial tool that eleven families of Kutunia Nua Sahi slum lying on private land, and therefore not entitled to land rights, may not be lying on private land at all. Whenever a revenue parcel boundary passes very close to a set of slum houses, then the benefit of the doubt should be given to slum dwellers, and a final verification should be done in the field. Field verification showed that those houses were indeed not on private land. Not only did this solve the land rights problem for these eleven families, which had been continuing for more than a year.

It thus also increased the efficiency of resource allocation by the revenue department by targeting field verification only to specific problem cases identified by the digital tool. This meant that teams had to be deployed only when necessary, thus substantially reducing unnecessary risks to government staff during the pandemic. But perhaps the most significant gain was the change of the attitude of the revenue department towards the use of digital data. Being cartographically oriented through their work, they embraced the new technology when they saw its ability to solve their technical bottlenecks.

5.2.4 Aiding the process of slum relocation as part of the JAGA Mission

- **Identification of slums on untenable land:** Slum relocation is a sensitive issue. The GoO had almost reached a deadlock, as spreadsheets extracted from GIS data from the JAGA Mission database showed that out of 1725 slums, approximately 50 percent of the slums would need relocation. However, the geospatial analysis through the tool, using the same data but with the geo-coordinates, showed otherwise. The logic was simple and shown visually. As individual land rights were being granted, it needs to be seen which houses were affected. If any slum is on some sort of untenable revenue land, the spreadsheet data showed all the houses located on untenable land, which may not be the case. Just because some houses are affected does not mean the whole slum is untenable. The tool demonstrated the power of spatial information to the Principal Secretary and showed that only 120 slums might need relocation (from 800-900 from the excel database). Further filtering and analysis would further decline to 90-95 slums, and these are slums that are totally on the railway track or exactly on a flood zone. This furthered the confidence in the tool.

- **Identification of land for relocation of slums**. Once the list of slum households located on untenable land was determined, the next stage in the process would be the tagging of the houses for consent to relocate. The ULBs identify tracts of land available for relocation of the slums, and they are always not very generous in releasing lands. Therefore, there is inadequate information available to check the land bank. Since CPR was working with the government in solving the issues of slum relocation, as part of the JAGA Mission, the government provided the data, bought from the Orissa Space Application Centre, which is a satellite-based survey done of all government vacant land for all 114 cities in Odisha. They understood the logic of the process and the need for spatial data to resolve this issue. This data and the analysis indicated more vacant lands than indicated by the ULBs. So, this data enabled the central level to negotiate with the local level and bring transparency to the process of slum relocation. So, the discussion was around the vacant patches pinpointed by their geographic coordinates, previously not identified by the ULBs, and challenges and resolution around the plausible vacant plots.

5.3 Size and volume of data

Post the pan state data generation as part of the JAGA Mission; the government had resorted to manual mode, as they could not use the GIS data. It was losing trust incapability of the data that they had created. The sheer size and the volume of the data also generated, in their way, acting as enablers. The government soon realized and became aware that the huge database they had generated and the scale at which they were working needed some structuring, organization, and management. The CPR partnership provided the technical expertise to do the needful.

5.4 Open Source provided flexibility

The choice of an open-source platform provided the flexibility to perform just in time geospatial analysis, or just visual representation of the GIS data, to address issues on the ground. In addition, a function could be developed to run the analysis. Finally, if the issue was to be replicated at scale, it was developed into a module that could be run anywhere from the web-based interface.

6 Challenges

6.1 Range of tools under development, but adoption uncertain

Currently, there are some tools in various phases of development for the GoO. For example, the Slum Land Information System (SLIM) is mainly a web-based platform that tracks the progress of slum land rights progress made on the field and the causes for the delay – a monitoring and evaluation tool. In addition, there are other such tools/platforms like the DIGIT Platform, Mo Sarkar App, etc., being developed by different consultants for various specific functions for the GoO. The GoO is aware of the City App Tool, and the partner CPR is working with the GoO and using it on their request to resolve issues. This tool intends to enable geospatial information and analysis in policy decision-making and implementation. Also, this is the only verifiable tool that is ready to be put up in the GoO servers. However, currently, there is uncertainty about the tools/platform that the government plans to adopt and make it part of its servers.

6.2 The varying architecture of the existing tools impedes collaboration

The tools under development may be in different stages of development and may have different architecture and platforms. For example, they may or may not be using open-source components, like the City App Tool, but might have a very different architecture. In addition, they may have used different software, and also, the models for the development of the software are different.

6.3 Data collection challenges

- One of the areas of concern is the time that is required for data generation. As mentioned earlier, for effective data usage, data creation needs to be pre-planned, pre-processed, and organized so that post-processing time is minimal and data quality is ensured. But, again, this needs time and resources, which are rarely acknowledged by government data collection efforts.
- Another aspect is that there are data overlaps in many organizations, and in some cases, data collection efforts are not synchronized with time and space.
- In some organisations and government ministries, data produced is closely guarded and siloed, allowing data sharing and use. This leads to duplication of data collection and sub-optimal use of resources and does not allow a comprehensive understanding of related issues.

6.4 The clarity in data use and utility of tool use

Data collection becomes the main focus without an adequate understanding of how data can be used sustainably. It is often driven by top-down compliance requirements or different project monitoring or planning needs. Ensuring incentives for data use will be a key lever in improving capacity to leverage and use data effectively. The tool demonstrated the use of the spatial data generated and clarified the potential use of the robust data for JAGA Mission implementation.

It is, of course, a challenge to develop the software creatively. Still, it is even more important to have great clarity on what the government wants and can do with the data, how one cleans the data, structures the data to enable efficient and useful decision making. How data influences policymaking remains unclear, making it difficult to understand which specific data points would be most useful. The clarity that the tool developer should have been able to clarify to the client the usage of the tool, what exactly the tool can offer, and more importantly, what is it the exclusivity of the tool- why should they adopt one tool over the other (compared to other competing tools).

6.5 The existing model of software development is not sustainable

Currently, government agencies' most popular model of such tool/dashboard development has been to engage a company with geospatial expertise for data collection and then another software development company to do the tool development. The software (mostly expensive, proprietary software) and the skills belonged to the private companies contracted to prepare the digital data sets. When the project contract is tendered out, with some terms of reference for developing it, it is handed over to another entity for operations and maintenance. Once the contract periods end, the digital data is handed over to the government, not the software or the operational skills. The process of development is not participatory, not collaborative. Even if using open-source components, the process followed is closer to the model of developing proprietary software. The tool is not developed internally to participate in government activities but is outsourced. Therefore, there is a need to engage with people who are deeply involved in the domain within the government system to understand the exact features of the tool so that it is useful.

However, the City App toolkit was developed collaboratively within the realm of implementation of the JAGA Mission, on an open-source platform, in collaboration with a scientist for algorithmic expertise, and other consultants as domain specialists about data context, government officials who know the technicalities of the revenue land and other on-ground realities, and other domain consultants within the Mission. As a result, the tool services the needs of the JAGA Mission and has the potential to build upon its robust, scalable architecture.

6.6 Required skills for geospatial analysis are rare

One of the key barriers to the use of data in the public sector is the lack of skills and incentives amongst data producers/users (e.g., data managers, revenue collectors, different departmental data, etc.) to continue to generate useful data and analyse, understand, and present the data collected into easily understandable messages that are relevant to decision-makers. Therefore, limited training of data producers and lack of incentives for data producers to understand data use and analysis is an important barrier for stakeholder capacity to use data.

There is a lot of deliberations and discussions around geospatial data and analysis. However, the required skills are not readily available. It is not about the availability of specific know-how of various programming languages or knowledge of geospatial technology. Still, the skill to creatively utilize programming knowledge with geospatial aspects and in-depth internalization of the context. The challenge stems from a lack of understanding of the technology itself, the internal sciences of the GIS software, the scientific parameters, and the interlinkages between to develop functions about resolving real issues on the ground. Very few people would have both geospatial knowledge and knowledge of computer and web-based programming, with the creativity of bringing all these aspects together, closely knit with the problems that need to be resolved.

The enablers and the lessons learnt through tool development provide the foundation for scaling up and replication in any context. The lessons learned are not unique to Odisha, and the JAGA Mission data and its application but are also key challenges to effective data usage, tool development, and usage of a tool, in general.

7. Lessons learnt for scaling up and replication.

7.1 Data collection needs to be focused on to minimize post-processing time

For the JAGA Mission, the extent of data collection was huge. Data was generated for 1725 slums and 1 70,000 households. Based on the Principal Secretary's suggestion, the questionnaire was kept very short and dedicated, delving mainly on basic household parameters that were needed to give the slum dwellers land rights. All unnecessary questions were avoided. The number of questions was kept at a minimal. The advantage was very clear because when a pan state survey of this scale is done, there are different data collection agencies with varying skills and facilities. Some may have a tablet; others might be doing a pen and paper survey and then transferring it to excel. So, lots and lots of spelling errors and other errors occur rampantly. It becomes a huge job to clean the data. However, if your columns are minimal and the questions are focused, it is easy to manage the messy data collection process. The survey process will be massive, with lots of errors, and post-processing is never easy unless pre-planned, monitored, and managed extremely well. The size and the focused questionnaire helped this massive data generation exercise complete within 6-7 months.

7.2 Importance of data structuring and organization

- **Conflicts with GIS and statistical databases:** Simple technical aspects are ignored in the planning stage before data collection is initiated. Thus, conflicts arise at a later stage of data collection, rendering some of the data useless in the long run. For example, the certain format of GIS files, say shapefiles that store vector map data would allow only certain characters maximum to fit into the column's header. However, the survey questionnaire may be developed, not taking this into cognisance because they are not aware of GIS formats and their limitations. So, when such basic issues are not attended to in the pre-planning stage for data generation, when responses come back, the column head may be truncated, and the data making no sense (if the question was- 'does the family have access to toilets" in the questionnaire, it gets translated to the column headings in the GIS database as "Does the f"). However, these kinds of technical issues can be managed during post-processing if the number of questions is minimal. For example, no other questions have the same ten-string characters. More importantly, there is someone within the team to have the domain knowledge to spot such issues.
- **Naming protocols**: Another example where the data was handed over in a format run on software would have 1725 layers having the same name. The submitted data was in different folders, having the names of different districts. Each district folder contained sub-folders of the different cities, and within them were the names of different slums. In folders, so there were just layer names with no designator. So, in Adivasi Sahi slum, the layer name was identical to the layer name in Anand Sahi Slum. So, there are 1725 layers with the same name. So as long as you keep it in the folders, you

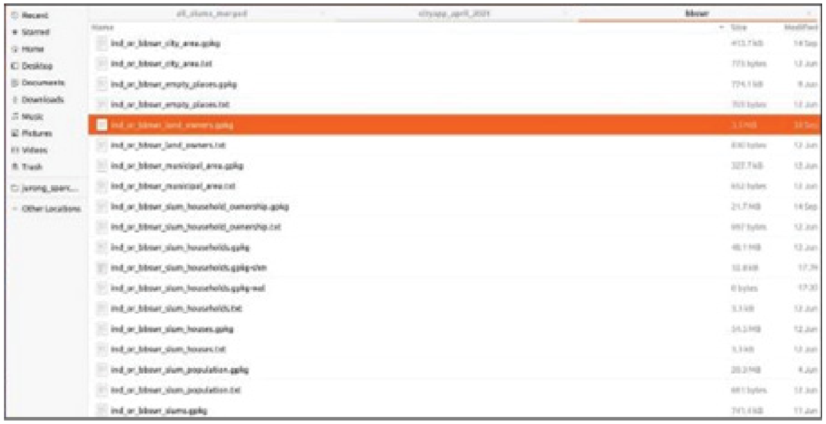


Figure 7: Organised data sets

Figure 8: Creation of metadata

know which layer belongs to which slum, but when you open it in software, all the layers all look the same. Sometimes the layer names or the column headings are indicated as some temporary codes and thus becomes impossible for someone other than the person who coded it to decipher. So, when consultants generate the data and are not associated with the government beyond the project duration, these issues become difficult and useful data cannot be extracted. Not following proper naming conventions leads to high post-processing times, ineffective data usage, and this is important learning that needs to be conveyed.

- **Metadata:** Another important aspect most Indian data sets do not contain is metadata (data about data). Census or National sample survey explains the metadata. But here, all the data created in these numerous projects that are in progress, there is no metadata. The tool enables the organisation and creation of the metadata. So anytime any attribute table gets updated, it is updated in the metadata text file. Each of the layers will maintain its metadata. The problems of Indian data gathering are not so much analytical ability later on but about managing the data at scale and organising people with diverse skillsets to follow a uniform method. Data pre-processing, pre-planning, and pre-organization are not usually addressed by the development community. Pre-processing cannot wait until the data generation exercise is completed; it must be addressed earlier in the process. This is critical for large data sets created for the large schemes in India.

7.3 Process for dynamic data collection

Data collection is steadily increasing. However, its usability is determined by the data's timeliness, credibility, accuracy, and completeness. Inconsistencies in these parameters make their use effective for decision-making. Several governments have introduced management information systems to streamline data collection and quality. While adequate physical and digital infrastructure is necessary for collecting, analysing, and sharing data, there is a need for sufficient communication between data producers and users. The disconnect between those who are collecting and producing the data and those who are going to use the data results in cases where the data collected is not fit for decision-making by users. Data collectors are not involved in identifying what data is needed. Overall, this leads to limited or inefficient data use. Thus, continuous engagements with data users and data producers, with participation from government program staff, partner organisations, and other key relevant domains through the data cycle, significantly impact the use of the data. The spatial data generation process (focused on the need), with CPR demonstrating the use of the tool, positioning within the government implementation ecosystem, testifies this fact.

7.4 Software development and knowledge transfer close to the area of application

The outsourcing model of software development is unlikely to work. The key to scaling up is keeping the software development close to the software application area, internally developing the functions as required. Close both in terms of geography, if possible, should be developed with the relevant stakeholders. CPR's steady presence and confirmation that technical contribution would go beyond project timelines and the trust built within the government departments have helped them advise them on many fronts beyond their initial scope. So now, if any digital project is underway, they would consult with the CPR team, try to analyse that system slowly, and eventually internalise the concepts. The learning process is getting internalised through regular engagement with the relevant stakeholders. Capacity is being built through on-the-job practice, a continuous process, and not limited by project timelines. Manuals and documentation of the data structures, organization of data, etc., can also supplement the process. The knowledge transfer is mutual between the external technical consultants and the in-house domain experts.

7.5 Identification of use cases

Seeing is believing. Application of the tool to real-life issues continuously attending to the dynamic nature of urban issues is key to sustaining the tool. It not only aids the government to own and manage the data to resolve critical spatial issues and provides agility and transparency to policymaking, thus optimising resources and enabling inter-departmental collaboration.

7.6 Scalability of software

The software needs to be flexible, adaptable to the ever-morphing complex urban policy issues that need spatial analysis and interventions. The software is built on a global grid, thus is already built for scale. Also, an open-source tool meant that one is not restricted by the existing state of any software, nor by what the proprietary software is deciding as to its modules/functions. The user dictates the software, and not the other way around, unlike proprietary software.

7.7 Invest in internal capacity

There are different ways capacity can be built into the system. Till 2020, 100 JAGA fellows were hired. These JAGA fellows were young professionals and worked as field agents. They were hired through Tata Trust and were embedded in the different municipalities. They played an active role in facilitating the initiative by mobilising communities to participate in ward-level meetings to finalise the slum upgradation work, participatory need assessment, monitoring, communication, and implementation of the JAGA Mission. When their term ended, most of them got absorbed into the government system in different municipalities, and they continued as arms of JAGA. During their tenure as JAGA fellows, they learned how to work in government projects and also contributed to enhancing government capacity. So, there was absorption in terms of human resource personnel and the JAGA fellows' enhanced skills. So, dependency on external capacity was averted in a way.

Investment in this type of model of building internal capacity and not outsourcing small day-to-day operations will be imperative. Investing in departmental geospatial capacity so that the personnel can add layers of data, conduct quick visualizations of the data, and operate some in-built modules, will be the first step to take this tool forward. External specialists and internal domain experts could advise and develop the modules as needed. Still, the onus would be on the government to operate the system as and when needed. This is the alternative to the outsourcing model, where knowledge transfer is usually incomplete. This will be critical in building the internal capacity of the government and empowering the departments in using geospatial data and analysis for their day-to-day activities, aiding evidence-based decision-making.

8. Way forward

The tool is envisioned to bring about ***spatial empowerment of ULBs*** and CBOs by allowing them to use and get command over their data. CPR is currently hand holding the GoO in familiarising with the tool. The capacity and the confidence in the tool's usefulness are gradually on the increase. CPR will need to continue to handhold till the capacity is created within the departments of the GoO. In terms of capacity building in operating the tool, the government officials (users) need familiarity with the software or how these digital technologies work, being well versed in cartography concepts.

The ***current state of urbanisation in Odisha poses an opportunity***. The urbanisation trend over the last five decades in Odisha reveals that the smaller towns (Class IV, V, and VI) are growing faster than the bigger towns. Also, these small towns are more in number and geographically spread over different parts of the state. In contrast, the larger cities are fewer in number and concentrated in specific advantageous locations, attracting more urban populations with higher opportunities. Therefore, the familiarization of the use of such tools in this state of urbanization can be envisaged as an opportunity to empower the small cities with the use of data, mainstream the use of such tools in their day–to–day work, and inculcate a culture of use of geospatial data and analysis for decentralized decision making, thus optimizing their scarce resources.

However, to collect, manage, organize, and effectively use the data, the geospatial capacity of the state and city administration needs to be enhanced through the inclusion of GIS operators at the district and city level (along with the current MIS experts) and Geospatial experts. Software developers are embedded in a full-fledged data/geospatial cell at the state level. The focus of capacity building would be to ***build the internal capacity*** of dictating the use of the tool, albeit with the use of external expertise and advisory as needed.

Currently, the tool is operating on detailed Dhenkanal data as a pilot. The tool should expand and include all available data in the state. The tool has been designed to be mainstreamed as part of government operating structures under the realm of the JAGA Mission. The vision of the tool is to ***aid inter-departmental coordination***, where different departments work together not only in sharing data but resolve issues transparently using this platform to visualize and analyse the data. To mainstream such a tool within the government structures, it would need further deeper and elaborate development with the government. Such a tool can be anchored in H&UDD and embedded in the OUHM or the AWAS housing program as part of the JAGA Mission.

Another aspect of the tool is its versatility. The tool has been designed so that separate query functions can be created and converted to a module, which can be treated as a separate subject of analysis altogether. Also, it is built on a global grid platform. It can also aid different domains of urban policymaking, being part of a domain agnostic geospatial laboratory.

CPR as a think tank can enhance its work on evidence-based policymaking by exploring and expanding use cases of such geospatial tools to other domains***. Creating a digital geospatial analysis lab*** could be an option where the City App can be placed and used for solving various geospatial problems in collaboration with a range of government departments. This will allow inter-disciplinary use and collaboration through dialogue and cooperation between different experts/departments to resolve complex issues that cut-across different departments/stakeholders. The resolution of the geospatial problem can then feed into policymaking. This can be hosted on CPR's web server. Funding for the above can be proposed to different donors or even the government, for specific use cases, to sustain and reap not only the powerful benefits of the tool for evidence-based, participatory policymaking but, more importantly, ***promoting transparency in the governance processes***.

Annexure: Online tool for city-level Slum free planning

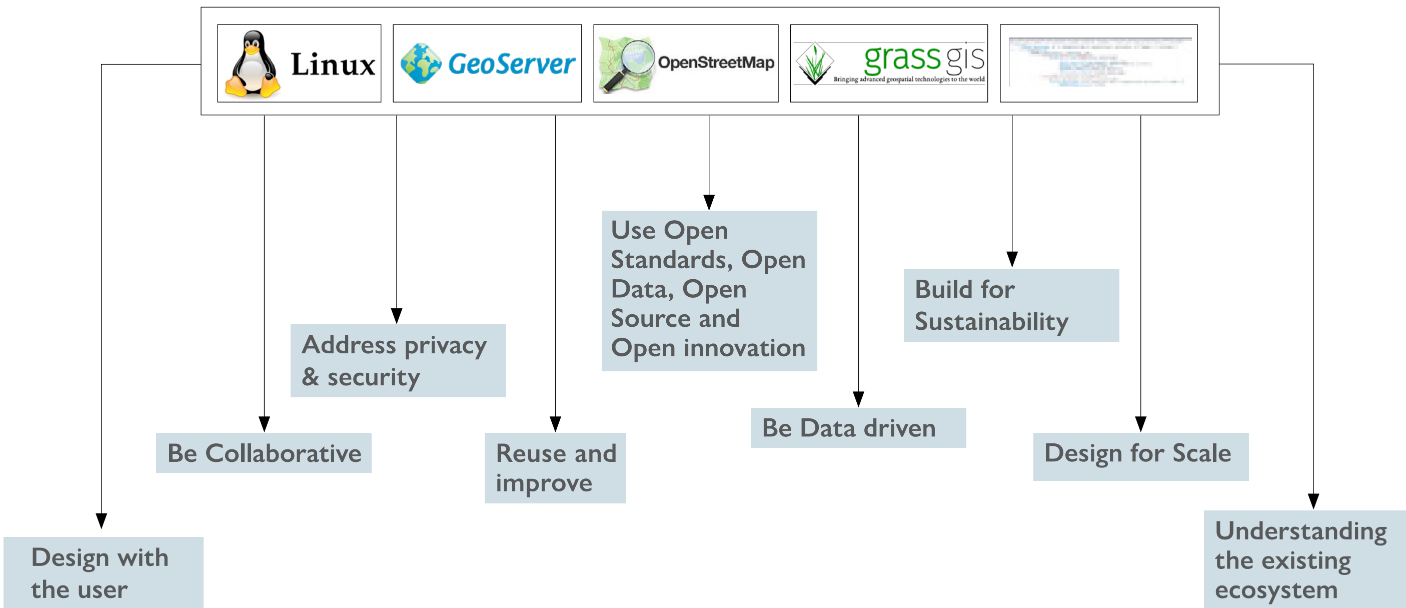


Figure 9: The tool corresponds to and strengthens the nine principles for digital development

The workflow of the online tool for city-level Slum free planning follows the logic of a regular GIS work session.

- **Key components**

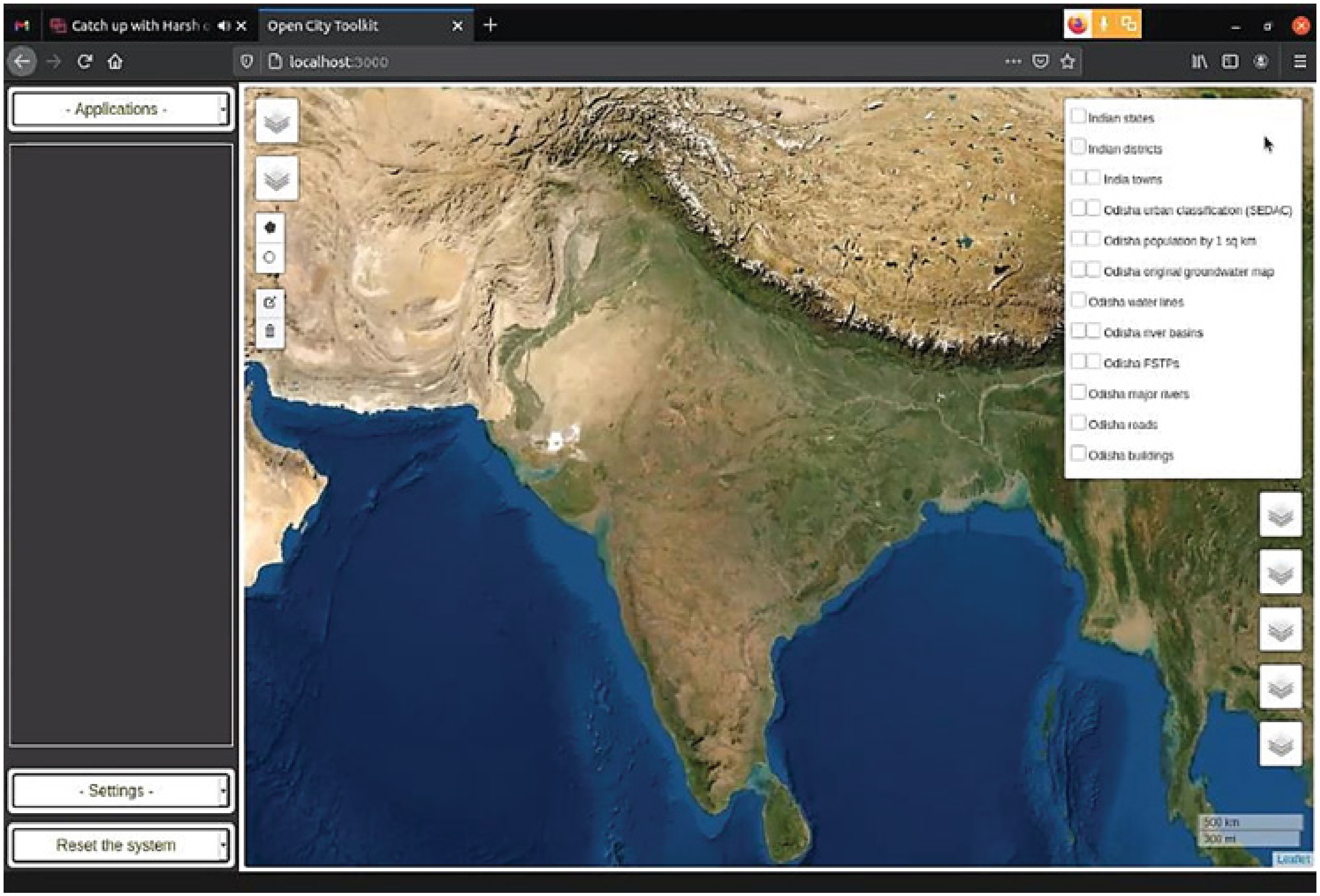


Figure 10: The tool user interface –concept[2]

[2] https://github.com/citysciencelab/open-city-toolkit/wiki/1.-Toolkit-for-Open-and-Sustainable-City-Planning-and-Analysis-%E2%80%90-User-manual

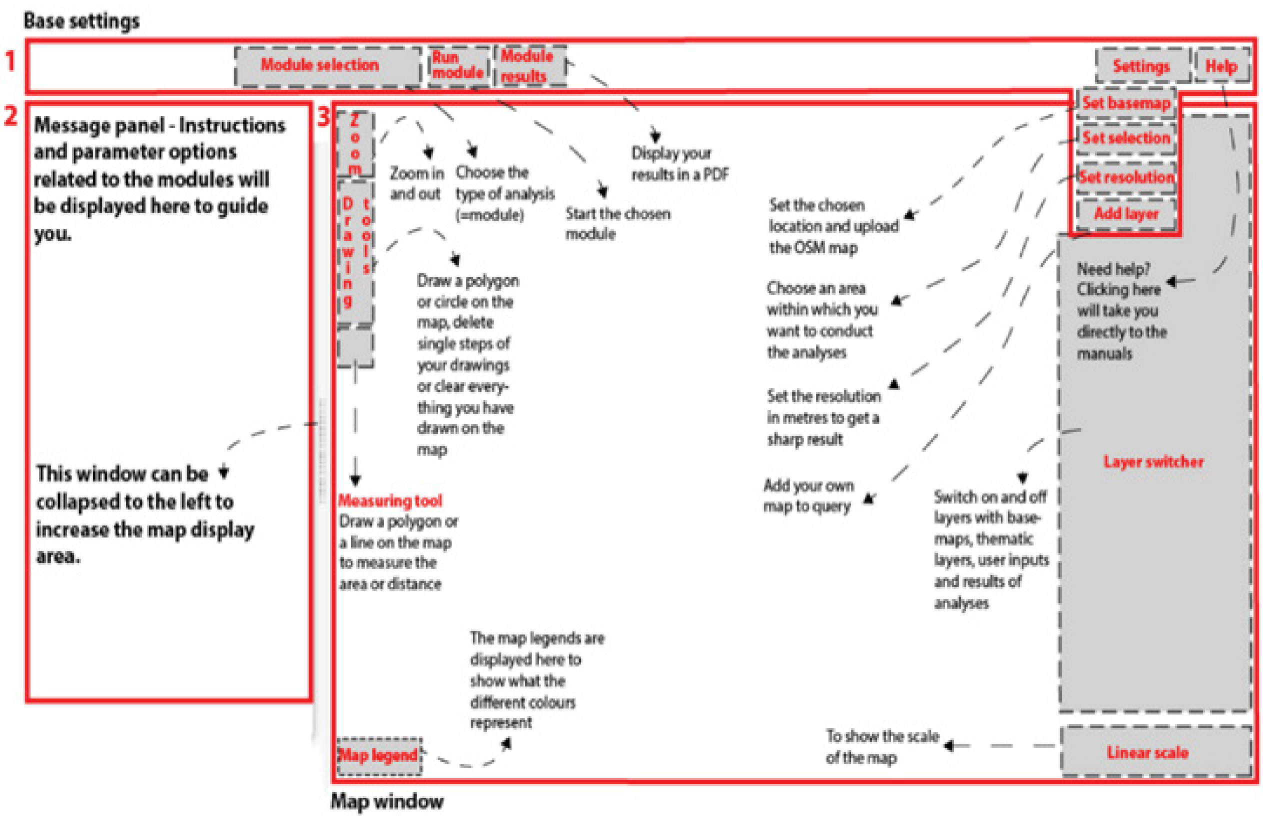


Figure 11: The tool user interface –settings[3]

1. **Base map:** The base map contains the basic map layers of the area of concern. Then, the user selects the concerned area of analysis /visualization. The base map- a set of map layers forms the basis of the spatial analysis. The base map contains the geographic area of concern's road networks, buildings, and water lines. These three map layers are extracted from an OpenStreetMap (OSM) file uploaded by the user. The base map is a general-purpose map.

2. **Map layer:** This layer is a geo-data file containing geographic features related to a particular topic. The user first uploads the relevant map layer/s into the tool to conduct analyses on certain topics. These map layers can be stored as part of the tool. GIS experts usually prepare the map layers based on the available data.

3. **Selection:** The user sets a selected area before conducting any operations, and calculations are performed only for this selection.

[3] https://github.com/citysciencelab/open-city-toolkit/wiki/1.-Toolkit-for-Open-and-Sustainable-City-Planning-and-Analysis-%E2%80%90-User-manual

The Settings menu has a drawing tool that helps the user demarcate any area. When the process is complete, the user can view the outputs of the analysis of the selected area.

4. **Resolution:** Spatial resolution is the precision at which raster-based analyses (e.g., isochrone calculation) are performed. Recommended resolution for analyses on a city scale, 5-10 meters, and for analyzing a very large area (200 km^2), is around 50 meters. A higher resolution setting implies less computing time, but the low quality of the analysis and the resultant PDF will have a lower definition of the resulting image.

5. **Modules:** Modules are computational functions that generate specific analytic results. They are the core features of the tool.

The architecture of the tool

The Open City Toolkit connects several external tools to implement a flexible and easy-to-use web GIS solution. A Linux system equipped with several applications is required as a base system. The modular framework for the open-source web-GIS platform is planned with distinct front-end (for web browser-based visualisation, discussion, querying) and back-end (data storing, GIS engine, algorithms, analysis). The flexible architecture is illustrated in Figure A12 and has the following tools:

1. **Geo Server:** An open-source server written in Java allows users to share, process, and edit geospatial data. Designed for interoperability, it publishes data from any major spatial data source using open standards[4].

2. **Geographical Resources Analysis Support System (GRASS) GIS** is open-source software that has been under development for more than 28 years, integrated with well-tested and documented algorithms into a joint GIS suite. Users and developers communicate through an online source code repository, mailing lists, and a Wiki to review existing code and develop new methods. GRASS GIS runs on common operating systems (MS-Windows, GNU/Linux, Mac OSX), giving basic and advanced functionality to casual and expert users.

3. **Node.js:** Node.js is an open-source server environment. Node.js allows you to run JavaScript on the server.

[4] https://en.wikipedia.org/wiki/GeoServer

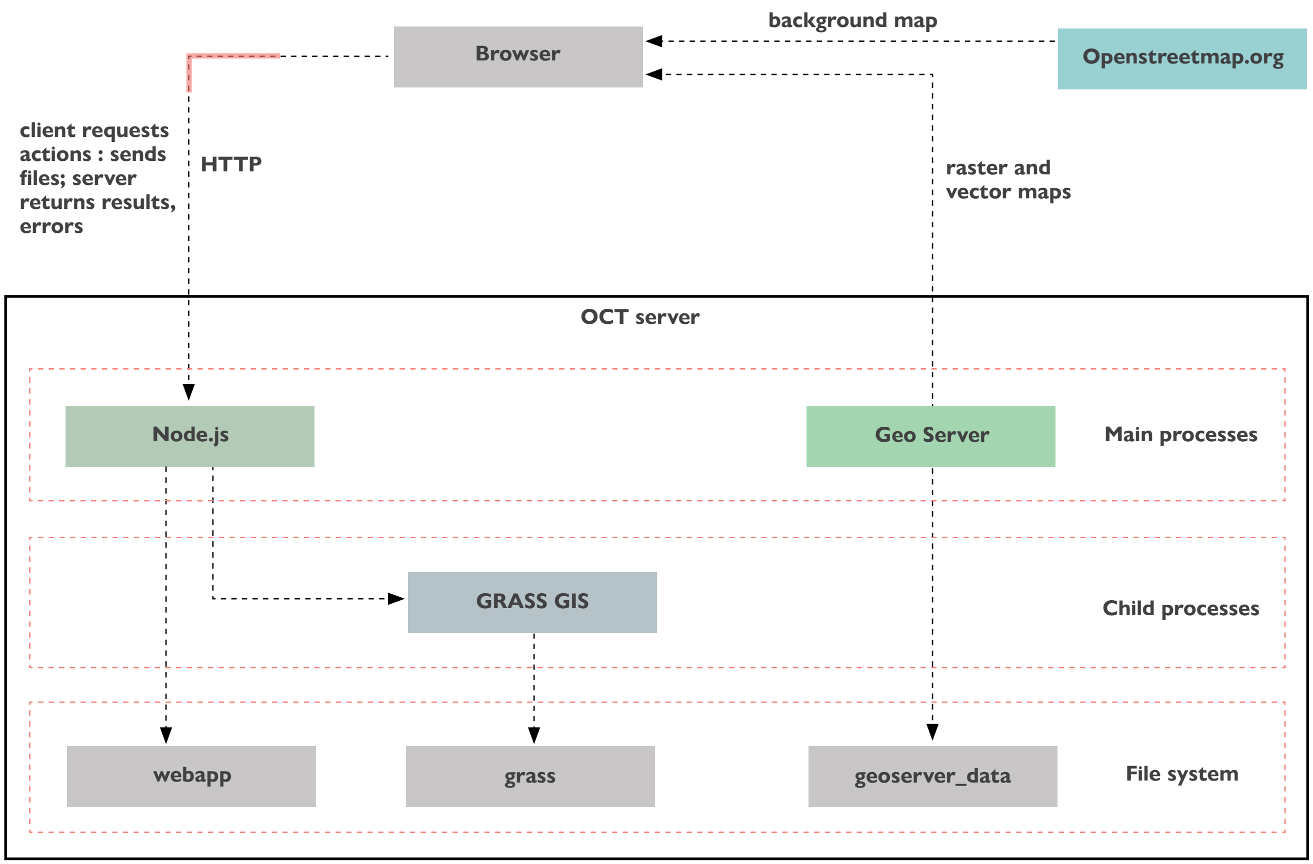


Figure 12: Architecture of the Open City Toolkit

Prepared under project titled "Building State Capacities for land, planning and slum upgrading for resilient cities" supported by Foreign, Commonwealth & Development Office (FCDO), Govt of UK

Authors

Shubhagato Dasgupta, Sujaya Rathi, Antarin Chakraborty, Baisakhi Sarkar Dhar

Acknowledgement

The authors are thankful to Principal Secretary, Housing and Urban Development Department (H&UDD), GoO for his constant guidance and feedback. The authors are also grateful to the officials of HUDD, GoO, Dhenkanal city team and the technical team of JAGA Mission, for their support, valuable time and information for preparing the knowledge product. The authors also acknowledge the FCDO team in India for their constant support and valuable inputs towards finalising the knowledge product.

Publication design

Trinankur Banerjee

E: trinankur@gmail.com