URBAN CONTEXT MODELLING FOR HUMAN SENSOR WEB

Habtom Tsega, Rob Lemmens, Juma Lungo, Menno Kraak

1 Department Of Geo-Information Processing, University of Twente, Enschede, The Netherlands
2 Department of Computer Science, University of Dar es Salaam, Dar es Salaam, Tanzania

ABSTRACT

Nowadays, widespread use of mobile devices became an opportunity for advocates of Participatory GIS (PGIS). People carrying affordable mobile phones can be utilized as moving sensory servers who can generate a lot of useful observation about the usage and prevailing situation of public services (such as water, health, sewerage, sanitation etc.) in their immediate vicinity. Theoretically, challenges of urban public services can be observed and reported by human sensors on a real-time basis. Thus, the concerned governing body can immediately act upon the reported problems. This is the governing principle behind the Human Sensor Web (HSW) System. HSW is an innovative notion of participatory sensing by ordinary citizens (also known as ‘human sensors’) with the overall target of improving the provision and sustainability of public services. Better yet, the utility and reliability of the content generated by citizen can be improved if the context of use of the system is carefully studied and modelled. The aim of proper handling and processing of context is to augment the ability of the HSW system to sense and act upon information about its immediate personal, social and physical environment. Given its wider applicability and possibility for sharing among systems, studies show that ontology-based modelling is the most suitable approach to represent context of use of an application. In this case, context is an input for use in conjunction with citizen generated content in attempt to create a context adapted geo-information suitable for decision makers and planners regarding their dealings with provision and management of public services. Accordingly, creation and maintenance of ontologies has become an engineering process that needs to be handled via guided approach. A central theme of this paper is to deliberate the practical ontology development process for the purpose of modeling the context of HSW. Ontology development process is an approach that dictates the construction of concepts and relationships meant to capture realities regarding the domain of interest at hand. Such an approach is rarely combined with methods recommended for context modelling. In our case, we attempted to adapt the Unified Process for ONtology development (UPOn) method in conjunction with an ontology based context modelling method called PIVON. We applied and tested the approach to build and maintain contextual knowledge base needed to populate our ontology model.

KEY WORDS
VGI, context modelling, ontology engineering, urban context, Human Sensor Web, UPOn, PIVOn, Unified Process
INTRODUCTION

The Human Sensor Web (HSW) is a recent development within the field of Participatory GIS and founded on the standard principles of Sensor Web; a standard set by OGC® to enable the integration of heterogeneous sensor webs into spatial data infrastructure (SDI). The notion was pioneered in 2009 as part of h2.0 initiative of Google.org and UN-Habitat [11]. It deals with the engagement of routine citizens as geosensors in generating Volunteered Geographic Information (VGI) [26]. Human Sensor is the name given to people equipped with affordable mobile phones and other ICT tools to sense their immediate surroundings and report their observation. By the same token, Human Sensor Web (HSW) describes the networking of human sensors and the use of content generated by them.

The sensing task can be pertaining to different geographic phenomena. Focusing on (urban) public services, it can be about such as status (e.g. water point misfunctionality, drug stockout in health facilities), disasters (e.g. flood, fire), incidents (e.g. crime, corruption), accidents (e.g. traffic) and the like. In this case, the overall goal is to enhance the usability and applicability of geo-referenced data/information by not necessarily specialist users [12] on the one hand, and to ensure the participation of citizens in sustaining public services on the other hand. This way, users of the services are enabled to generate spatial content regarding the prevailing situation of the service with the help of affordable ICT alternatives. More importantly, information produced before or after the analysis of data gathered from voluntary citizens are disseminated and made accessible to interested endusers via mobile and Geoweb interfaces [34].

The benefit of citizen generated content can be improved by improving our ability to interpret it. However, HSW operates in different contexts. For instance, what is considered as ‘human sensor’ (who reports) might differ from context to context. Interpretation of a report generated by a human sensor in a given context is totally different from another context. This makes interpretation of the data generated by a HSW system depend on the context it operates. If we consider an example of water quality, what is ‘pure water’ in one context might not be the same in another context. Therefore, interpretation of HSW data is inaccurate unless the context of its use is properly understood. Knowledge of the context is crucial to semantically interpret and geo-map the reported messages, to make space-time analysis in order to extract beneficial information out of the messages (such as regarding authenticity of messages), and finally disseminate the information in a format more suited to the situation of end-users (e.g. to decision makers). We cannot achieve those goals without the proper understanding of the knowledge of the context in which human sensors generated the spatial data and endusers prefer to visualize the processed information.

Context is the knowledge about the social, physical and temporal situations describing the usage of HSW from human sensor and enduser perspectives. According to the commonly applied definition by Dey and Abowd [1], context is any information about an entity that helps to characterize its situation and behavior. An entity is a person, location or object considered relevant in the interaction between the user and an application. Example of HSW context as applied in the water domain includes,
location of water points, proximity between water points, mobile network coverage, language preference of information presented as text, type of water source, water users and vendors etc.

In order to make context-based interpretation and analysis of data, context knowledge has to be captured and modelled. Due to its wider applicability and reusability, ontology-based context modelling is the most preferred technique. In this paper, a generic method is adopted to guide the process of converting informal context reality into its formal representation via ontological principles. The purpose of the modelling process is to create a conceptual ontology model that holds knowledge about the context of HSW. The purpose of the ontology in turn is to capture the context knowledge and become an input in the analysis of VGI contributed by human sensors. The HSW that uses the context ontology in order to analyze the VGI submitted by human sensors is considered as Smart HSW. It provides improved benefits to decision makers and governors by fulfilling their demand for contextually relevant information. Particular to urban context, it gives a better control of the information flow for handling tasks in urban planning processes and decision making for improved governance, decision making, and service improvement.

CITIZEN AS SENSORS

Today, a number of personal, GPS-enhanced, and mobile technologies [Phones, bicycles, laptops, palmtops, and wearable devices] and participatory Web (Web 2.0) have the capacity to collect, disseminate and analyze geo-tagged data in real-time manner. HSW is founded on the utilization of such technologies to make experts and non-experts alike voluntarily contribute to the creation and use of spatial data. One of the potential application areas of HSW is in community-led services provision and monitoring by transforming volunteered humans (i.e. ordinary citizens with portable devices) into geo-referenced sensory servers [12]. For example; in urban water services sector, HSW can be considered as a PGIS tool to facilitate the efficient two-way interactions between urban water service givers and takers. The prototype implemented by Zanzibar Water Authority (ZAWA4) of Tanzania for the purpose of community-driven monitoring of water availability and sanitation is a practical example in this case.

There is a spatial dimension to the information usable in policy and decision making with regard to urban public service delivery and monitoring. The participation of citizens in this process is justifiable for two main reasons [29]. First, for the sake of efficiency and sustainability, it is crucial to give attention to what people living around the service delivery centers have to say about the status of the services they receive. This converts citizens to sensors who contribute to the generation of spatial data about the prevailing status of the urban public service infrastructures. Secondly, the use of spatial information by non-experts and the novice makes citizens become part of the potential users of information outputs resulting from the analysis of service-related spatial data generated by themselves.

Mobile technology is most widely and commonly used comparative to other ways of communication [2, 28, 18]. Given this widespread penetration and delegation of mobile phones including into novice users, mobile-based applications (called mApps) provide great opportunities for governmental and non-governmental organizations to reach out and interact with citizens who should not necessarily possess any sophisticated skill [14].

http://geonetwork.itc.nl/zanzibar/
Among advantages of collecting citizen data (via HSW) is that it is the best option in areas where there is time and cost constraints to build large networks of material sensors. This makes HSW affordable by all countries including the poor ones in the southern hemisphere; because it requires a minimal infrastructure [16]. Additionally, HSW can rely on cheap and commonly used communication technologies [such as Short Message Service - SMS]. However, human sensors are not limited to mobile device and SMS use only. Other ICT devices like Personal Digital Assistants (PDAs), Personal Computers (PCs) and other communication protocols like Multimedia Messaging Service (MMS), the Internet, and so on are also possible. In principle, volunteered individual citizens or social user groups that represent them can send their observations pertaining to real world events via available multimodal (short text, pictures, audio and video) messaging system [23].

As an exemplary tool of Participatory GIS for the gathering and use of VGI, the HSW system helps to facilitate contribution and role of the local citizens [26] in demand-oriented policymaking and continuous monitoring of the efficiency of public services using spatial knowledge. There is a relatively greater potential of utilizing HSW application in urban setting as a generator of VGI for tackling urban problems (such as regarding water service functionality monitoring). This is because mobile and Internet users in urban environments significantly outnumber rural mobile users [17]. In this case, HSW enables urban dwellers to become moving sensors who observe, sense and contribute to the generation of spatial information pertaining to urban problems. By so doing, urban human sensors are voluntarily involved in the creation of Spatial Data Infrastructure (SDI) that helps in solving the reported problems (e.g. broken water pumps) in particular and urban public service planning and monitoring in general. To that end, the overall target is improvement of urban public services.

CONTEXT AND CONTEXT AWARENESS

Definition of context itself depends on the ‘context’ in which it is viewed. According to Cambridge Advanced Learner’s dictionary [5], context is “a situation within which something exists or happens, and that can help explain it”. The applicability of context in computational and non-computational sciences varies significantly. For instance, Nash et al. [24] referred political context as “those aspects of the world that are relevant to policy actions, it is the arena of those actions”. This includes aspects such as the distribution of power, the range of organizations involved and their interests, and the formal and informal rules that govern the interactions among different players in the political ground. For development actors seeking to influence policy, political context matters because it determines the feasibility, appropriateness and effectiveness of their actions.

In the field of Human Computer Interaction (HCI), usage and definition of context focus on two different views – user-centered context view and system-centered context view [3]. From the user perspective, context is relevant information useful to characterize the situation of a user interacting with a system. It can be regarded as information about where the user is, who the user is with, and what resources are nearby [27]. On the other hand, context can be viewed from a system perspective. Meaning, context is regarded as any information that the system senses and acts upon in order to make constant adaptation to changing situations.

Both views are important in our definition of context. We consider context as a hybrid of user-view and system-view perspectives. It is conceptualization of different relationships between the
representations of the situation of the user and the application as well as other relevant things evolving around them. We seek to describe the key human and non-human actors and their interrelationships in the operation of HSW platform in a preferred domain. The purpose is the realization of a context-aware HSW platform – which we designated as Smart HSW (see section 3). Context awareness is the ability of a system to perform reasoning about the context and trigger the necessary actions to adapt to the changing situations [10]. For this purpose, context knowledge needs to be gathered and represented in a format accessible by the Smart HSW system that uses it. This process is known as Context modeling.

To formally model and utilize context as machine-readable input for information processing, many context modeling approaches have been in practice. Strang et. al. [30] in their work surveyed and evaluated different types of context modelling approaches. Due to its strong features; namely, expressive power, hierarchical organization, formality, standards basis, support for programming abstraction and interoperability, their evaluation favored ontology-based modelling. Ontology is a formal description of a conceptualization and is becoming the standard practice for conceptualizing and formalizing context knowledge. It enables the creation of a shared understanding and knowledge base about the entities and relations in a given domain. It is reusable and can be easily shared across systems. Thus, we prefer our context model to be based on ontological principles. For reasons of wider applicability, reusability, and adaptability, many other scientific studies [10, 35, 36, 5, 6, 20] affirm the preference of ontology for context modeling.

[fig. 1]: Top-level context taxonomy and their semantic interconnection – An Actor and a Device are the two entities involved in HSW. An Actor carries a Device or a Device is carried by Actor. Entity is located in a given Environment (e.g. location) and uses a Service (e.g. information visualization). The service is displayed on a Device

In this paper, we recommended the adoption of an ontology-based modeling approach referred to as the Pervasive Information Visualization Ontology (PIVOn) [15]. The rationale is because it is found to simplify the task of context conceptualization by identifying top-level classifiers for which ontologies would be formulated and organized - user, device, service and environment. [fig. 1] illustrates the
abstract categorization of context and the semantic connection between the classifiers. We redefined ‘user’ into ‘actor’ to make it include human sensors (who perform mobile sensing) and end-users (final users of processed information). We consider that these four classifiers embody the overall context of HSW and represent the context scope and sources outlined in section 4. To conceptualize the context of each classifier, the method acclaims the formulation of a matrix whose answers will eventually help to populate the context model. It does so by corresponding the four context elements against five ‘WH’ questions (What, Who, Where, When, Why) as seen in table 1. To provide realistic responses to each inquiry in the matrix, the PIVOn method is utilized in combination with the UPON ontology development practice discussed in section 6. The combination is crucial in order to acquire realistic representation of the context knowledge about the use of HSW in the chosen domain. In short, the context ontology process tries to give answers to each relevant question in the matrix and make a formal representation of those answers with the help of ontology.

<table>
<thead>
<tr>
<th>Classes</th>
<th>WHAT</th>
<th>WHO</th>
<th>WHERE</th>
<th>WHEN</th>
<th>WHY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actor (A)</td>
<td>What actors &amp; doing what</td>
<td>profile &amp; social relations</td>
<td>Where actor is &amp; performs</td>
<td>When actor performs</td>
<td>Why actor performs at</td>
</tr>
<tr>
<td>Device (D)</td>
<td>What devices used</td>
<td>Who use the devices</td>
<td>Where devices are placed</td>
<td>When devices are available</td>
<td>Why devices are included</td>
</tr>
<tr>
<td>Environment (M)</td>
<td>What environment &amp; objects</td>
<td>What actors</td>
<td>Where objects are placed</td>
<td>When objects are available</td>
<td>Why environment organized in this way</td>
</tr>
<tr>
<td>Service (V)</td>
<td>What Web service</td>
<td>Who uses services</td>
<td>Where services are offered &amp; used</td>
<td>When services are offered</td>
<td>Why service offered &amp; used</td>
</tr>
</tbody>
</table>

Table 1: 2D matrix of context taxonomy

THE USAGE OF CONTEXT ONTOLOGY

The content and utility of ontology is extremely diverse. The purpose of our ontology is to formally represent the context of use of HSW and make it accessible to the analysis of citizen generated data for better interpretation and presentation of the content. Thus, it can be referred to as Context ontology. The HSW notion states that citizens contribute to the generation of spatial data and relevant information are made accessible to relevant audiences. But this statement is highly simplified because it disregards several practical issues some of which are addressed by proper management of context knowledge through ontology.

The main assumption behind this work is the belief that HSW performs better if it is made aware of and adaptable to context than a contextless HSW. Through prior HSW piloting [34], it became apparent that context unawareness affects the capability to interpret and use the wealth of data crowdsourced from citizens. Such a system is merely contextless; meaning it does not take context into account and it is unaware of the situation in which citizens generate observational data and endusers consume processed information. Our interest in context is pushed by the need to effectively exploit the spatial information generated by HSW. Those benefits are demonstrated below by describing the key features of a HSW which is made smart with the help of knowledge about context.
The motivational questions remains to be “why do we need the HSW system to be aware of context?”. In other words, what will contextual HSW have which contextless HSW doesn’t? or how will context add value to HSW?

SMART HSW

Our definition of smart HSW assumes that a HSW system gets smarter with the utility of context than without. It is capable of utilizing context knowledge when performing analysis of HSW data and it fits information to the situation of an intended decision maker. This definition holds three features that a smart HSW should fulfill by means of its awareness and adaptively to contextual knowledge. Those features are discussed below.

**HSW that Listens correctly**

Precise listening and understanding what the human sensor ‘tells’ is one of the challenges of HSW implementation. This is not just a matter of interface as applied in Human Computer Interaction (HCI) [19]. It involves a great deal of semantic challenge because humans and computer machines interpret information differently. The HSW requires a technique more advanced than human-to-human communication using natural language or machine-to-machine communication using standard xml syntaxing. The HSW should be equipped with the capability to transform human-readable version of reports (written in natural language) into their proper computer representations.

The initial concern is how sensors should ‘speak’ (input their observation) to make the system ‘listen’ and interpret semantically accurate message (as intended by the human sensor). Today, smartphones, mobile phones have image capturing, audio recording, gps-based location identification, and tracking capabilities. The use of devices other than mobile phones (such as personal computers) is also an advantage. Focusing on the most commonly used texting functionality by smart and simple phones alike, human sensors can register alphanumeric data concerning the issue they are observing. This textual data has only implicit semantic meaning attached to it. Before analyzing and mapping the message, the system has to have a better control and interpretation of what is what. For instance, questions like ‘which of the messages are about water point functionality problem?’ has to be understood and sorted out correctly.

To establish smooth understanding between the human sensor and the SMS-based HSW system, the quickest solution seems to predetermine short codes that would easily be extracted by the system. However, practical experiences [11] show that this technique highly affects the usability of the system because human sensors tend to prefer the use of natural language i.e. use the habitual way of texting than sending codes. In this case, to interpret the free unstructured text, the system requires a contextual knowledge describing the situation of the generated message. Correct understanding of the message is the initial requirement before imagining how citizen generated data can be further used and presented. To do so, knowledge about the context can be utilized to enhance the listening ability of the HSW system.

For instance, imagine that the HSW system received several texts telling about the situation of water point functionality. Some may describe the case as broken, or non-functional water point. Some may tend to put the time in absolute (e.g. Monday) or relative (yesterday) terms. Some may specify the location (e.g. place name) or location information might be missing so that it should be derived
through other indirect methods. The first three basic questions to interpret the coming messages are "what is the report about?", "when did the reported case happened?", and "where did the case happen?". From the registered free texts, these and similar questions can never be properly answered without the knowledge about the user, time, device, location and similar contextual factors.

**HSW that analyzes smartly**

Once the meaning and semantics behind data entered by citizen are understood with the help of contextual knowledge, context could also be of further help to enhance the use of the crowdsourced data. Aided by Geovisualization and spatio-temporal analytical techniques; the system can be made capable of contrasting content against context in order to extract new forms of information. For instance, an urban planner might need to be informed about the trend of the reported problem over time. What is causing the problem might have some seasonal factors. This requires thorough analysis of the citizen generated data with the help of temporal and locational context knowledge. The purpose is to detect expected and to discover unexpected information [22] by means of visual and interactive tools. Geovisual analytics is a set of methods of geospatial data presentation and analysis through the use of exploratory, highly interactive and visually enabled information production tasks [22].

**HSW that knows what to present to whom**

It is important to account the relevance of information to different parties according to their particular context. This is based on the principle of context-based geovisualization [9]. Every type of information is not equally important and relevant for every types of known user. For instance, aggregated and summarized information make sense at the higher level of the government hierarchy. On the other hand, information about a single broken pump might suffice for a water vendor who live next to the water point and whose life depends on it.

Questions like “What type of information, what presentation method and to whom?” are highly dependent on the context knowledge. For instance, point map that shows where the non-functional water points exist are of no help to users of simple mobile phones (because they do not support maps). In that case, information in the form of text can be a coping mechanism. Still, language could be another constraining contextual factor because, for instance, message composed in English is useless if forwarded to someone who does not understand the language. If a water user has to be notified about the existing status of a water point, it makes sense if locational context is taken into account, i.e. such information would only be helpful to people who live in close proximity to the water point under consideration.

**LEARNING CONTEXT**

The modelling of context intends to create a formalized context model, i.e. a digital representation of the context under which HSW operates. One of the priority questions is 'which context is to be modelled?’. In most cases, this context is not distinctively and explicitly recognizable. It is inherently and implicitly hidden within targeted applications and domain. To detect those contextual elements and make their representation concrete and formal, it is important to follow certain guidelines (discussed in section 6). However, before explicating context, we must demarcate the context within
which HSW operates; which we technically called the context scope. For the sake of brevity, clear distinction was made between two sources of implicit context information within the context scope for HSW – application context and domain context.

Application context is a collection of facts describing the functionality of citizen sensing systems. Mainly, it is about by use of devices and services in the Smart HSW that uses the context ontology (i.e. Device and Service context classifiers in [fig. 1]). It is learned from selected HSW-like applications which are already operationalizing or have been operational in the past. They should not necessary be about the targeted domain (e.g. water) as long as they exemplify the principles of HSW and PGIS. For instance, in Tanzania, where our Smart HSW is expected to be deployed, the following applications are relevant – the first HSW implementation in Zanzibar (Zanzibar-HSW\(^5\)), mobile reporting system for human rights violation (by Commission of Human Rights and Good Governance), mobile application for corruption incidents (m-rushwa), mobile-based reporting of stockouts of malaria drugs by health workers in public health facilities (SMS for Life) and the like.

By studying pioneer HSW-like apps, functional requirements of the software that uses the context ontology are determined. Those requirements form the foundation of the application context. For instance, context information about who is the human sensor may depend on the choice of the persons or social groupings that the system regards as senders of observational data. Such information are uncovered through the understanding of actors and their roles in the domain of interest; which makes up the domain context. Application context is incomplete without the domain context which describes the characteristic of the preferred domain of interest. In our water point functionality example, types of water users or water service providers, administrative hierarchy, economic or literacy status of users, network connectivity or coverage etc are parts of the domain context. The Actor and Environment context classifiers in [fig. 1] stand for domain context.

The first task in the whole process of context modeling is learning and understanding the context to be modelled, a task which we labeled context cognizance or learning context. The process involves set of different activities that should be followed to become familiar with the context under consideration by admitting feedback of knowledgeable experts. The process enables the detection of implicit contextual parameters and evolve them into their explicit digital representation. However, the detail and relevance of the context information depends on what we want to do with the context model – what sort of context analysis is needed. For instance, we should consider ‘proximity of human sensors to the geographic phenomenon they report about’ as relevant context information only if we intend to use the context model to check the credibility of messages and if we consider ‘proximity’ as one parameter.

In this paper, the modeling process deals with the particular context of HSW for reporting water functionality in urban Tanzania. The context of urban Tanzania for citizen sensing has been studied through qualitative research approach. It involves series of learning activities and systematic gathering of context information onto which investigation and explicating of contextual factors depend. Those activities mainly constitute cyclic interviews and discussions with relevant experts and professionals who are capable of explaining about the targeted applications and domain. Those experts are sources of information on application and domain knowledges – they are called Domain Experts (DEs). In the

\(^5\) [http://geonetwork.itc.nl/zanzibar/](http://geonetwork.itc.nl/zanzibar/)
case of urban water point functionality, DEs are people who know the relationship between water users and providers in Tanzania as well as how HSW-like softwares works under the principle of VGI through the use of mobile phones operate. Relevant information gathered in this manner gives raise to the building of a knowledge base that characterizes the application and domain context of our Smart HSW.

Expert participation is an important aspect of the context learning process for context modelling (for detail see section 6). Feedback and consensus of experts is needed at every stage of the process to assure if the contextual facts and concepts extracted at different stages are representations of the realities in the context scope. Initially, selected officials and specialists in the preferred domain are asked basic questions regarding structure of management information system and functionality of tools used. Knowledge about the application logic of citizen sensing systems can be acquired by involving users, developers and hosts of existing such systems. The questions include; Who are the main users? How information flows? How information is analyzed? How information is presented? Those questions are derivations of the WH-questions tabulated in table 1. In the step-by-step conversion of implicit context hidden in the collected unstructured information into their explicit representation, interim results are continuously presented to similar experts in order to check if the semantic knowledge encapsulated in the context concepts matches the reality. Those tasks are supported by document reviews, self-learning, participatory observation, and guided tutor of targeted applications and prior knowledge about the domain.

MAKING SENSE OF LEARNT CONTEXT

The end result of the context modeling process is an ontological representation of the targeted context. It constitutes domain ontology (DO) and application ontology (AO). This separation of concerns is required because the sources of context are not identical; as defined in the previous section. The former sub-ontologies is a representation of the application context and the latter sub-ontology is a representation of the domain context. Otherwise, no sharp boundary can be drawn between the two ontologies. The context model intermingles the two sub-ontologies because the model is an abstraction of the use of the Smart HSW system in the chosen domain.

Application Ontology (AO) constitutes concepts describing the functionality of the human sensing system. It answers questions like; who reports, what type and form of messages, what type of device, what is done with citizen generated information and to who is processed information targeted to and the like. Construction of AO requires understanding of how citizen sensing systems work in practice. For this purpose, similar systems that are already running in the context scope need to be studied.

Another substantial body of the context model is the DO. It holds context information regarding the domain under consideration. In our case, the use of HSW system in the water domain is targeted; particularly for reporting problems related to water point functionality. For eg. human sensor reports can be about broken water pumps. This makes the DO to be information about the relationship between users and providers of urban water service in Tanzania. For this purpose, officials in the ministry of water, at district level, water users and water vendors are considered to be the main sources of knowledge about the domain context.

To illustrate how learning of context takes place, let us assume a simple example. Imagine a Smart HSW that receives reports about a broken water pump by a chairperson of a water committee. Based
on the context classifiers identified in section 2, The chairperson (Actor) is in charge of a water committee (Actor); which makes ‘chairmanship’ the relationship between the two actors. It is also apparent that the chairperson is a type of human sensor (Actor) who reports (Service) to Smart HSW regarding the functionality of water pump (Geographic phenomena with location or Environmental context). There are other implicit context ideas that can be inferred with the additional knowledge of the domain – water pump is a type of water source and has geographic space or human sensor reports from a specific location at a specific time. In this example, ‘noun’ terms like water committee, water pump, location, time, functionality etc hold context information in them and are termed as lexicons. Other ‘verb’ terms that interconnect lexicons (such as reports, located, receives etc) are regarded as relationships.

**MODELLING STRATEGY**

The context modeling process requires a choice of an approach with the ability to perform activities of learning context from the identified sources and transfer the learnt context into formalized representation in the easiest possible way. The modeling strategy should follow an approach that properly integrates application and domain contexts discussed in the previous section.

One option of integrating the application and domain contexts is through a process called ontology merging [31]. This means, the development of separate ontologies for each source of context (AOs and DOs), and later merge those small ontologies to construct the bigger ontology in order to make a holistic representation of the context. This strategy calls merging to take place at final stage of the context modeling process, i.e. merging the formalized ontologies constructed for each context source. However, integration of two or more ontologies is complicated, labor-intensive and error prone [31] and should be performed if the ontology cannot be constructed otherwise. However, existing ontological resources are considered as assets in ontology development process and can be integrated into the context model through this mechanism.

Our preferred strategy is to make the integration of context from the variable context sources at the initial stages of the modeling process. This means, before the context parameters emerge into a formal representation, the integration can take place during the extraction of context lexicons which are the basis of ontological concepts. The UPON method (see section 6) is recommended for transforming informal context information into formal context ontology concepts by treating the integration at lexicon level; i.e. when lexicons are extracted from the collection of context-level information gathered from different sources. The UPON method has lexicon extraction steps for application context as application lexicons (AL) and domain context as domain lexicons (DL), and later make merging between the two sets of lexicons to form Reference lexicons (RL). The RL is a list of lexicons that holds a representation of both application lexicons and domain lexicons. Those are the lexicons that will grow into a more formalized context concepts; thus in way the final formal context ontology is an integration of AO and DO.

**ONTOLOGY ENGINEERING: FROM IMPLICIT CONTEXT TO FORMAL REPRESENTATION**

The ontology development process should be guided by a well-founded methodology. There is no universally accepted method for developing ontologies. Brusa et al. [4] identified two groups of ontology development methodologies. On the one hand, some methods are based on specific project
experiences [13] or enterprise models [33]. On the other hand, there are more generic methods proposed to accord with the widely used principles of software development process. The methods that fall in this category are highly preferred when requirements are not clearly understood at the beginning [4]. Methods like METHONTOLOGY [21, 7] commend evolutionary refinement of ontology through activities that take place throughout the life cycle of the prototype of the software that utilizes it. Others such as the 101 Method [25] suggest iterative and incremental development of ontology by pursuing the notion of starting small and growing bigger in the proceeding consecutive rounds.

Similarly, the Unified Process for ONtology (UPON) [8] is a software development approach for ontology building. There are comparative advantages of UPON, which made it the preferred approach to guide our ontology building process. It inherits the Unified Process (UP), one of the most widely used methods by the software development community. It capitalizes the extensive experience accumulated for decades within the software engineering arena. It also upholds the essence of evolutionary, incremental and iterative development of ontology by continuously validating interim results of different activities in the process. As a use-case driven approach, it simplifies the task of acquiring knowledge by identifying actors, roles and relationship in the domain and application.

![Context Ontology Development framework](image)

[fig. 2] The Context Ontology Development framework (adopted from [15,8])

Also, the UPON approach clearly defines the involvement of the sources and engineers of Knowledge. People who are knowledgeable about the domain of interest and selected applications are treated as Domain Experts (DEs) and their role is to feed information about target applications and domain of interest needed as a foundation for later extraction of context ontology concepts. They also validate interim results produced during the process of ontology development by comparing them to their real-world meanings. The role of the Knowledge Engineers (DEs) is to facilitate the collection of those expert information, and to later organize and formalize them guided by the ontology engineering activities of the UPON. It also integrates the contexts of the application and domain by systematically merging the application and domain lexicons before the formalization process begins.
PROPOSED CONTEXT ONTOLOGY DEVELOPMENT APPROACH

The principal concern in our process of ontology construction is the systematic translation of informal context knowledge about the domain of concern into conceptual formal concepts of ontology. Commonly, the practice of ontology development process involves a combination of methods. There are a number of studies on ontology-based context modelling principles and approaches, however they lack a well-defined technique that properly takes account of practical contextual knowledge gathering, organization and formalization.

In our case, the requirement is a systematic gathering of knowledge regarding the context of the given application domain and to make sense of the collected knowledge in such a way that it gradually develops into formalized ontology representation. For this reason, we propose an empirical approach for context ontology building; by combining UPON [8] ontology development method (see section 6) and the PIVOn [15] context modelling approach (see section 2). This combination was justified because both methods complement each other. The PIVON method simplifies the context modelling task by creating a guideline which composes basic WH questions to facilitate the collection of basic context information and a structure for formalizing and populating the ontology model. Also, the UPON method makes the context modelling task founded on empirical knowledge collected from the context scope and has a possibility to include feedback of experienced experts. It facilitates the conversion of unstructured information into formal context ontology model. [fig. 2] depicts our context ontology development framework.

The framework constitutes five ordered building blocks – Lexicons, Glossary, Conceptual Semantic Network, Formal Ontology (written in standard Web Ontology Language, OWL) and Refined Ontology. It is indicated by blue skewed curves in [fig. 2]. These building blocks can be considered as the interim results of the ontology development process. They range from completely informal knowledge about a domain and Smart HSW to a formal machine-readable version of the domain ontology. Lexicons are vocabularies extracted from context information as described in section 4. Glossaries are concrete definitions of the lexicons as applied in the given context. Conceptual network is formed by specifying semantic relationships that exist between context lexicons. This network is converted into an ontology model coded using a formal standard language called OWL (Web Ontology Language). After refinement through improved requirements, it is this model that is used by the Smart HSW.

The building blocks are made up of four dimensions – workflows, iterations, phases, and cycles. Each building block is an output of each of five workflow (requirements identification, analysis, design, implementation and test); respectively. See the workflows listed to the left of [fig 2]. A round of sequential execution of the six workflows entails one full iteration. The process must allow as much iterations as possible in order to revise and amend the building blocks (see the downward arrow to the left of [fig. 2]). The process pursues not only iterative but also a start-small-and-grow-bigger approach. This means, the requirements of the context ontology and the Smart HSW software that uses it gets refined as the process progresses. Therefore, a smaller version of the ontology is constructed during the first iteration and the ontology gets continuously improved until the associated HSW software is fully developed.

Each building block has width and height. The width is an indication of a cycle which includes one or more phases; i.e. Inception, Elaboration, Construction and Translation. The extent of the broken
horizontal lines in [fig. 2] shows the range of phases covered by each workflow. The height of each building block indicates the relative degree of task involved in a workflow in the corresponding phase. The phases range from simple brainstorming and preparatory aspects of inception, through the expansion of initial ideas with the help of actual domain/application knowledge and technical aspects of constructing context representation to translating conceptual model by comparing it with real-world realities.

Another aspect of the UPON method is its clear division of tasks between DEs and KEs during different workflows. The rightmost compartment of [fig. 2] illustrates the involvement in the form of line graph. The role and contribution by DEs is higher during the commencing workflows which progressively diminish towards the subsequent workflows replaced by role of KEs. This is due to the fact that first-hand knowledge about the application and domain is of high demand during the initial activities of the first workflows. As the process progresses, the activities require more and more technical expertise and computational skills by KEs in order to make sense of the informal knowledge contributed by DEs.

Our proposed context ontology development approach has been tested by conducting extensive fieldwork in Dar es Salaam, Tanzania. It was applied for the development of context model usable by a Smart HSW system in water domain – particularly water point functionality. The outcome is a machine-readable representation of the context of Water Point functionality in urban Tanzania. Further refinement of the ontology development continues in different iterations until the associated smart HSW gets fully developed following the improvement of the software requirements. Demonstrating all activities conducted in the process of context ontology development in the field is beyond the scope of this paper. The full description of the fieldwork experience can be found in [32].

CONCLUSIONS

This piece of work is part of a research project named 'contextualization of HSW’. This project strives to prove the assumption that HSW can perform better with context than without. We consider HSW that uses context as Smart HSW. It admits context as important source of knowledge in the processing of citizen generated data. It is capable of extracting contextually relevant insights by analyzing citizen data (e.g. trustworthiness of reports submitted by human sensors) and appropriately presenting the processed information (e.g. no map outputs for simple mobile phone users). But first, relevant context knowledge must be studied and modeled in a format compatible with the HSW software that uses it; a process called context modelling.

This particular work focuses on the process of context modeling which is the initial phase in the whole process of contextualizing HSW. Modelling context means context has to be formally represented and inputted in conjunction with citizen generated volunteered geo-information (VGI) for further analysis. Ontology is a formal description of a conceptualization and is becoming the standard practice for conceptualizing and formalizing context knowledge. It enables the creation of a knowledge base that is reusable and can be easily shared across systems. Thus, we prefer our context model to be based on ontological principles.

We adopted a method for learning and modeling the context of HSW for water point functionality in Tanzania. To make a realistic representation of context, fieldwork activities were conducted in Tanzania in order to learn and understand the context under which HSW operates. The activities
attempted to answer the question "What are the socio-physical context factors of (non-) human sensors that currently exist implicitly in the domain of interest and relevant software applications? And how to explicate them?".

To do so, we gathered context facts (mainly through extensive interviews and document reviews) about the domain of interest (i.e. water point functionality) and HSW-like applications that are currently functioning or have functioned in the past. The collected information uncovers important contextual facts such as the nature of the potential HSW actors, their relationships, the physical environment around them, devices used, system functionalities and so on. We adopted the practices of software development approach to transform the informal and implicit set of information collected from diverse sources into their explicit and formal representation and ontology concepts (with the consensus of domain experts). A machine-readable model populated with concepts that stand for contextual facts is delivered at the end of this context modelling process. The functional requirements of the HSW software that utilizes the context model might continuously change. Thus, the fine-tuning of the context model continues until the HSW software that uses it is fully implemented.

In summary, we recommend a generic and adaptable context modelling approach for HSW powered by ontology and developed through unified process. We implemented and tested the proposed approach by attempting to transform informal context knowledge gathered from the field into its formal ontological representation. This way, implicit socio-physical context factors of (non-) human sensors that currently exist in the domain of interest and relevant software applications were explicated. The context modelling process outputted ontology model ready to be read by the HSW that is envisaged uses it in order to demonstrate how context augments the capability of analyzing of VGI data and presenting the analyzed information.

REFERENCES


[31] Stumme G and Maedche A [2001]. Ontology merging for federated ontologies on the semantic


