**FloodFeed: An Ontology-Based Data Feed for Flood Sensor Knowledge Integration**

Mohd. Zuhaili Mohd. Rodzi¹, Nor Hidayati Zakaria¹ and Mohammad Nazir Ahmad¹

¹Faculty of Computing, Universiti Teknologi Malaysia

**ABSTRACT**

Flood problems in Malaysia have recently attracted much attention and prompted the search for solutions. Existing ICT-based solutions such as flood websites and systems mostly are operated in silos, which makes it difficult to achieve knowledge integration. Utilising pervasive computing technology such as the “Internet of Things”, FloodFeed is introduced in this paper as a mechanism to gather flood sensor data collectively among Internet users. The feed is inspired by the well-known Rich Site Summary (RSS) Feed constructed in XML format governed by upper layer ontology to provide semantic data for systematic knowledge integration. By harnessing the potential of knowledge integration, new solutions can emerge with the potential to overcome flood management problems.

**Keywords:** flood management, flood sensor, knowledge integration, ontology.

1 **INTRODUCTION**

In Malaysia, frequent flood occurrences have become a major issue due to rapid urbanisation and climate change. Although Malaysia is not directly affected by serious natural disasters such as earthquake, hurricanes, typhoon, tornadoes, tsunami and volcanic eruption, the country is not spared from flood disaster. Flood is the most significant disaster in Malaysia, with heavy rainfall experienced throughout the year. Apart from providing an abundant amount of water, a certain pattern of unusual rainfall for a long duration of time can cause flooding to occur.

Technological advancement in the ICT industry has brought the modern world new methods and approaches in knowledge sharing, integration and dissemination. The Internet has become a necessary tool in the domain of knowledge. In terms of flood disaster management, knowledge related to flood is currently disseminated through publically accessed websites (Katuk et al., 2009) which are accessible through mobile devices and are capable of reaching a wider audience compared to the traditional media. Recently, the exploration of the potential to merge real-world hardware and virtual systems has made remote device control and data acquisition possible; this concept is widely known as the “Internet of Things” (Kopetz, 2011; Atzori et al., 2010). Certain websites have integrated this concept by, for example, the direct video streaming of flood-prone areas which are monitored remotely through Internet protocol cameras (e.g. http://publicinfobanjir.water.gov.my/).

Although there are many flood websites in Malaysia that provide various types of information, such as meteorological data, road closures, water level measurement, rain gauges, health risks, volunteer efforts and many other types of information, these websites operate on a silo basis which defines the information and knowledge they provide according to their own understanding (Othman et al., 2013). The inconsistencies in information thus create confusion and pose difficulties for knowledge sharing and integration. Furthermore, as can be evaluated from visitor counts, most websites fail to offer the flexibility and interactivity that would attract the interest of the public to learn more about flood as well as to share knowledge with others.

Therefore, inspired by the capabilities of the Rich Site Summary (RSS) Feed to increase visitor traffic to a website and the potential of the Internet of Things concept, this paper proposes FloodFeed as a solution. To increase public interest and awareness of online flood information, the FloodFeed mechanism collates integration flood sensor data from various sources through XML-based syntax and establishes a common repository of sensor data, thus allowing the flexibility for the development of various web applications.

To overcome the inconsistencies in concepts, the backbone of FloodFeed is an upper layer ontology that highlights the essential processes required for knowledge integration (KI). The ontology is to be made common for all knowledge contributors. This establishes the semantics of the vocabulary and concepts which are represented in XML as one of the methods of implementation.

FloodFeed adopts the current trend of existing ICT operating platforms that have captured the market, such as Android, iOS, Joomla, Drupal and Wordpress, by providing flexibility that allows users to harness their creativity in pushing the platform to its maximum potential through the development of millions of applications and widgets around the world. This approach secures the sustainability of the technology for the long term and creates a sense of attachment and belonging in the users.
This paper is organised as follows. Section 2 introduces the ontology that is the backbone of FloodFeed and is developed as a common structure for knowledge sharing and integration. Section 3 explains the flood sensor as a prototype developed for proof of concept. Section 4 elaborates on how the flood data from the sensor are organised in FloodFeed XML format to enable KI, and provides an example of a website to demonstrate how FloodFeed integrates data from various sources. Section 5 discusses the challenges and potential directions in future research. The paper is concluded in Section 6.

II ONTOLOGY FOR FLOOD SENSOR INTEGRATION

Despite being introduced decades ago in various domains, especially the business sector, the KI concept is still implicit in nature and suffers from inconsistencies in interpretation that lead to various forms of implementation with different understandings and terminologies (Oritz, 2003). This creates confusion about what exactly defines KI and how it should work in any knowledge management initiative. Extracting the essential processes of KI determines the core processes that are common and compulsory for all KI initiatives. This paper proposes the concept of upper layer ontology from the perdurant and endurant perspectives.

A. Perdurant Ontology

In the ontological concept, a perdurant is something that is composed of temporal parts. Therefore, whenever a perdurant is present it doesn’t mean that all its temporal parts will be present. A perdurant can refer to events and processes, such as a rescue mission, race, conversation or business process. The perdurant ontology of KI basically identifies the essential processes in integrating the available and contributing knowledge that is supposedly common for all knowledge initiatives. There are many approaches to developing a perdurant ontology, and this paper favours the approach taken by Colomb et al. (2010), Colomb (2012) and Colomb and Ahmad (2007) which utilises the essential system model proposed by Dietz (1999). The essential system model focuses on object processes at the upper layer ontology in order to develop a stable and generic ontology as the standard for cross-agency interoperability.

Extracting the Essential Processes of KI. Extracting the essential processes of KI involves a number of steps. A review of the extant literature indicates that there are many theories and concepts related to KI from various perspectives. It is important to explicitly highlight the common perspectives in the literature in order to represent the essential processes of KI. Among the known foundation theories is the knowledge creation theory introduced by Nonaka et al. (2000) which elaborates on the concept of knowledge from the root level as something personal and internal developed through an understanding of the environment, work routines and the ability to deliver a particular job requirement within an organisation. The tacit nature of knowledge is made explicit in a codified form that later becomes a medium for knowledge transfer. The socialisation-externalisation-combination-internalisation (SECI) framework basically deploys techniques for preserving employee skills and talents. This is carried out through a continuous process of socialisation between employees, the externalisation of knowledge, the combination of current knowledge with existing knowledge, and the internalisation of knowledge so as to become a tool for better job performance.

The SECI framework presented by Nonaka et al. (2000) complements Grant’s (1996) idea of KI and the knowledge-based firm, as knowledge cannot be integrated without first being created. Grant viewed KI within an organisation as integrating the knowledge of the experts who exist among the employees. This theory places greater emphasis on knowledge application rather than just knowledge creation. This emphasis can be seen in Grant’s proposed methods, namely, rules and directives, sequencing, routines, and group problem-solving and decision-making.

Knowing such differences exist, and derived from an analysis of the SECI framework, Grant's theory of KI and practical outcomes in flood management, this paper proposes the following essential processes: identification, creation, assimilation, and evaluation. It is anticipated that these processes would preserve the essence of KI, while at the same time broadening the variation of end applications.

Identification: Knowledge needs a place from where the process of integration can start. Identifying the root of the knowledge nature will determine the compatibility of the knowledge to be integrated with other knowledge in the same family and knowledge field. Accordingly, this will enable the process of creation.

Creation: Once identified, the knowledge creation process occurs by comparing existing knowledge with past knowledge (codified experiences) and the outcomes become the basis of determining whether new knowledge can be created. This new knowledge would need to meet the requirements of KI, the core objectives of which are the solution of problems and the improvisation of knowledge.

Assimilation: From the creation process, knowledge is then disseminated among knowledge users. At this stage, the newly-created knowledge is expected to be assimilated with the knowledge users. This will
ideally enable them to utilise the knowledge with a certain level of expectation based on the purpose and initial need for its creation.

**Evaluation:** In order to measure the deliverables and effectiveness of the recently-created and assimilated knowledge, it is evaluated thoroughly based on the actual outcomes and feedback from knowledge users. The knowledge thus obtained becomes integrated knowledge upon fulfilling the expectations of knowledge creation.

**Essential Model of KI.** The notion of the essential model of KI proposed by Dietz (1999) stems from the problem that arises in the business analysis or requirement engineering phase in the systems development cycle whereby, in modelling an organisation, some informational actions can be separated from the essential business actions. Dietz (2003) later developed modelling approaches known as the design and engineering methodology for organisations (DEMO) method. DEMO was developed as a methodology to extract the essential business processes from their realisation, thus establishing an essential model. Although DEMO is mostly utilised in business organisations, the present research explores the potential of the DEMO approach to be applied in the knowledge domain specifically for the purpose of flood management.

As DEMO was meant for business organisations, mapping DEMO into a KI essential model for the flood management domain could be expected to produce some minor differences; the goal of the present study is to undertake the mapping and still preserve the notion of "essential". A brief description of DEMO can explain the elements involved in establishing the KI essential model. Basically, DEMO regards an organisation as having individuals or subjects that perform production acts (P-acts) and coordination acts (C-acts). By performing a P-act, subjects will achieve the organisational mission and objectives. Within a P-act, there exist C-acts that coordinate the execution of the P-act. Concerning P-acts and C-acts, three layers of abstraction are established, known as the business system (B-system), informational system (I-system) and documental system (D-system). P-acts and C-acts also appear to occur repetitively in a generic manner and these are referred to as transactions. During a transaction there is an initiator who initiates the request for the transaction, and an executor who receives the request and performs the P-act.

The starting point of DEMO is the B-system where original facts are created before the I-system and D-system come into action. Therefore, a complete model of the B-system is also known as the essential model of an organisation. There are four integrated models that constitute the B-system, namely, the construction model, operation model, state model and process model.

Although the DEMO method is much more extensive and comprehensive than described above, the information provided is sufficient to explain the development of the construction model of KI which serves as the basis for the remaining models as presented in Figure 1. In Figure 1, there are four transactions centred at T1, T2, T3 and T4 that each represent an essential KI process, namely, identification, creation, assimilation and evaluation. S1 is the executor for T1 where the identification process of flood knowledge according to the related fields occurs. It then initiates the creation process that is executed by S2 within the KI creation pool. The created knowledge is then disseminated to the flood-related management and response team which evaluates the effectiveness of the created knowledge. One extension of the KI construction model would be the involvement of other related knowledge elements and critical success factors in the knowledge framework, which is not covered in this paper.

**Ontological Representation of KI.** From the construction model of KI, a perdurant ontology is developed to represent the KI essential processes. It is expected that a perdurant ontology which resides at the upper layer ontology would be mostly stable and invariant for interoperating activities. Colomb and Ahmad (2010) argued that even though most ontology research involves ontologies with heavy reliance on data models, upper ontology research recognises that there are actions in this world beside data. In their research, Colomb and Ahmad represented Dietz’s (2003) P-act as a perdurant ontology and the C-act as an endurant ontology. Adopting their approach, the present study proposes the perdurant ontology for KI as presented in Figure 2.

**B. Endurant Ontology**

An endurant is something that is wholly present and does not have temporal parts, such as a person, a house and, in this study, the sensor flood data. For the endurant ontology in this study, the approach proposed by Guizzardi and Wagner (2004) for business modelling is preferred. Some endurant ontology concepts are borrowed and customised to meet the requirements of the flood sensor KI as follows:

- **<<Type>>:** An entity that has an extension and intention (can be determined through subtype and instance)
- **<<SubType>>:** Categorisation of a type
- **<<RoleType>>:** The role and purpose of a type
- **<<Phase>>:** Scenario by which a type is applicable
<<Moment Individual>>: An individual that is dependent on another individual
<<Relator>>: Independent individual that relates to a dependent individual
<<Quality>>: The element of measurement to meet a certain expectation or objective.

The combination of the perdurant and endurant ontology establishes the backbone ontology for FloodFeed. The proposed ontology is depicted in Figure 3.

III FLOOD SENSOR DEVELOPMENT
For the purpose of demonstrating the concept of FloodFeed, a simple flood sensor is developed which consists of an ultrasonic sensor to measure water levels and a temperature sensor. Both sensors are interfaced to a microcontroller for data acquisition and transmit wirelessly to a remote receiver. The ultrasonic sensor output is pre-calibrated in centimeters by an on-board microcontroller and the temperature sensor measures the ambient temperature in degrees Celsius. The hardware block diagram is depicted in Figure 4.

IV XML-BASED FLOOD DATA
Utilising the above proposed ontology, the data acquired by the flood sensor are structured and converted into XML format using JAVA-based software with custom tags. In addition to the water level and temperature sensors, the sensor location data (latitude and longitude) are included for mapping which can be acquired in real time with a GPS or preprogrammed into the microcontroller.

For demonstration and proof of concept purposes, the following XML tag and format are used:

```xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"/>
<FloodFeed>
  <location name="sensor2" lat="3.41625" lon="113.8154">
    <waterlevel>15</waterlevel>
    <temp>37</temp>
  </location>
</FloodFeed>
```

A. FloodFeed KI Website Example
FloodFeed data can appear as raw data or converted to useful information and applications. Assuming the data can come from various sources with a similar XML feed format, an example website is created to process the XML tag using Javascript and PHP Hypertext Processor (PHP) in order to extract each sensor data. The data are then combined and displayed in a readable format for online public access. A screenshot of the website is shown in Figure 5.
Figure 4. Flood sensor hardware block diagram

Figure 5. Screenshot of FloodFeed example website

Figure 3. Backbone ontology for FloodFeed comprising perdurant and endurant ontology
V DISCUSSION
The concept of FloodFeed delivers semantic data for inter-domain KI. It enables sensor data to be contributed by individuals or commercial organisations as public data. From here, various applications can be developed to cultivate interest with the potential to further explore the flood domain as a collective effort utilising the latest technology creating new knowledge as well as improvising the existing knowledge. From the initial prototype, several advantages and challenges are identified which can become the basis for further improvements. These advantages and challenges are explained as follows:

1) The upper layer ontology based on processes highlights the essential processes required for KI to be generic and systematic for long-term sustainability.
2) The accuracy of the contributed sensor data requires a mechanism to ensure the highest level of integrity.
3) Although FloodFeed is inspired by the well-known RSS Feed, the process requires a certain level of technical skill and can be further simplified to increase public involvement and awareness.
4) Adoption of the latest technology will improve overall effectiveness in the delivery of solutions to flood management problems which is the motivation of FloodFeed.

A. Implications for Practitioners
The ontology layer greatly benefits practitioners, especially software developers, as it provides a “blueprint” for establishing a common language for KI. Regarding implementation, the targeted audience includes members of the general public who are involved in flood-related research, such as owners of personal weather stations who can share their sensor data. The audience also includes IT officers and website developers in flood management-related government and non-government agencies. The upper layer ontology and FloodFeed concepts can also be explored by researchers in the field of KI for further improvements.

VI CONCLUSION
In general, FloodFeed integrates the sensor data through a similar XML format governed by an upper layer ontology for KI. With such foundations, FloodFeed is expected to become a platform for sustainable and effective KI to further innovate new solutions and improve existing methods to address flood management problems in Malaysia.

VII ACKNOWLEDGEMENT
The authors wish to thank the Ministry of Higher Education, Malaysia for funding this study under the Long Term Research Grant Scheme (LRGS/b-u/2012/UUM/Teknologi Komunikasi dan Informasi).

REFERENCES
Colomb, R. M. (2012). Representation of action is a primary requirement, “Ontologies for interoperating information systems”. In M. N. Ahmad, R. M. Colomb, and M. S. Abdullah (Eds.), Ontology-based Applications for Enterprise Systems and Knowledge Management, IGI, 68-76.
Kopetz, H. (2011). Internet of things. In $R e a l$-Time $S y s t e m s$ (pp. 307-323). Springer US.

Knowledge Management International Conference (KMICe) 2014, 12 – 15 August 2014, Malaysia
http://www.kmice.cms.net.my/